



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121831

TO: Jon Ashen
Location: REM-2B02/2C18
Art Unit: 1635
Friday, May 14, 2004

Case Serial Number: 09/673994

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Ashen,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 10:04:10 ; Search time 1380 Seconds
(without alignments)
11842.622 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 gggaggtgtgactttattc.....gaaaaaaaaaaaaaaaaa 3847

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3836	99.7	3847	3	AAZ35044
2	2628.4	68.3	2651	3	AAZ35053
3	1988.4	51.7	2336	3	AAZ35052
4	1671.8	43.5	1696	7	ACC83475
5	1504.4	39.1	1860	7	ACC83474
6	1498	38.9	1575	5	ABA82987
7	1392	36.2	1855	7	ACC83473
8	1367.6	35.5	1996	3	AAA93131
9	1220.2	31.7	1488	4	AAI59942
10	1195	31.1	9266	4	AAI59942
11	1195	31.1	9266	4	AAI59942
12	1195	31.1	9266	4	AAI59942
13	1195	31.1	9266	4	AAI59942
14	1115.2	29.0	1456	7	ADA53558
15	815.2	21.2	1558	4	AAI58156
16	815.2	21.2	1558	4	AAI58156
17	615.8	16.0	1824	2	AAI59942
18	576.8	15.0	581	3	AAZ35050
19	528.4	13.7	851	2	AAI59942
20	511.4	13.3	559	3	AAH97867
21	498	12.9	509	3	AAZ35045
22	485.8	12.6	500	3	AAZ35046
23	484.4	12.6	562	5	AAH97866

24	472.8	12.3	487	3	AAZ35047
25	450.2	11.7	586	3	AAZ35051
26	444.8	11.6	6271	4	AAK85395
27	444.8	11.6	6271	4	AAK85395
28	444.8	11.6	6271	4	AAK85395
29	444.8	11.6	6271	4	AAK85395
30	444.8	11.6	6271	4	AAK85395
31	444.8	11.6	6271	4	AAK85395
32	444.8	11.6	6271	4	AAK85395
33	444.8	11.6	6271	4	AAK85395
34	444.8	11.6	6271	4	AAK85395
35	444.8	11.6	6271	4	AAK85395
36	444.8	11.6	6271	4	AAK85395
37	444.8	11.6	6271	4	AAK85395
38	444.8	11.6	6271	4	AAK85395
39	444.8	11.6	6271	4	AAK85395
40	444.8	11.6	6271	4	AAK85395
41	444.8	11.6	6271	4	AAK85395
42	444.8	11.6	6271	4	AAK85395
43	444.8	11.6	6271	4	AAK85395
44	444.8	11.6	6271	4	AAK85395
45	444.8	11.6	6271	4	AAK85395

ALIGNMENTS

RESULT 1
AAZ35044
ID AAZ35044 standard; cDNA; 3847 BP.
XX AAZ35044;
AC AAZ35044;
DT 28-FEB-2000 (first entry)
XX Mouse CNREB-2 cDNA clone KND42.
DE
XX CNREB-2; CNRE binding factor; transcription factor; mouse; renin;
KW negative regulatory element; cardiovascular disease; hypertension;
KW congestive heart failure; diagnosis; therapy; ss.
XX Mus musculus.
XX Key Location/Qualifiers
FH CDS 1198..3489
FT /*tag= a
FT polyA_signal 3809..3814
FT /*tag= b
XX WO9955343-A1.
XX 04-NOV-1999.
XX 23-APR-1999; 99WO-US008502.
XX 24-APR-1998; 98US-0082997P.
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX Chen YE, Horiuchi M, Dzau VJ, Tamura K;
XX WPI; 2000-062002/05.
XX P-PSDB; AAY32375.
XX A new regulatory element binding factor used to treat an adverse
XX cardiovascular condition.
XX Claim 1; Page 77-79; 57pp; English.
XX This is the nucleotide sequence of cDNA clone KND42 encoding mouse
XX negative regulatory element binding transcription factor CNREB-2. KND42
XX was obtained using yeast one-hybrid cloning and mouse kidney cDNA library
XX screening. CNREB-2 binds to the 5' flanking region of the renin gene and

modulates expression of the gene. 2 Other isoforms (see AAY32376-77) of CNREB-2 have been identified, the shortest isoform showing strongest binding activity to the CNRE of the renin promoter. CNREB-2 shares no significant homology to any other known gene or protein. It is a novel zinc finger transcription factor. Expression vectors, isolated nucleic acid molecules and CNREB-2 polypeptides are claimed. The invention also relates to CNREB-1 polynucleotides (see AAZ35040-43), polypeptides (see AAY32374) and inhibitors, and their use in claimed methods for decreasing or increasing renin expression, and for determining the level of CNREB-1 expression, an increase in expression being indicative of an increased susceptibility to development of a renin-angiotensin system mediated disorder

XX Sequence 3847 BP; 805 A; 1008 C; 1171 G; 852 T; 0 U; 11 Other;

Query Match 99.7%; Score 3836; DB 3; Length 3847;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAGGTGTAGCTTTATTCATCAAGATCAGTCAACGTGATGAAGAAACGGGAGCTTCG	60
DB	1	GGAGGTGTAGCTTTATTCATCAAGATCAGTCAACGTGATGAAGAAACGGGAGCTTCG	60
QY	61	GTATTGTCTCATCCATTGTGGAAGAGTGTAGTCTAAAGTTGGGAGGAGNGNTGTNTA	120
DB	61	GTATTGTCTCATCCATTGTGGAAGAGTGTAGTCTAAAGTTGGGAGGAGNGNTGTNTA	120
QY	121	AAACAGANAATTTTTTTTGAGATGTGACGACAGCTTGGGAGTTAGATTCACTGAANACAA	180
DB	121	AAACAGANAATTTTTTTTGAGATGTGACGACAGCTTGGGAGTTAGATTCACTGAANACAA	180
QY	181	TTAGTGACTTTGATTAAGANGNGGTNTTCTTCAATTTCCAGATGTAGCTGGGCTTC	240
DB	181	TTAGTGACTTTGATTAAGANGNGGTNTTCTTCAATTTCCAGATGTAGCTGGGCTTC	240
QY	241	AGAAATNACNGTTTTTTTTTTTGATGATGATGATGATGATGATGATGATGATGATGATG	300
DB	241	AGAAATNACNGTTTTTTTTTTTGATGATGATGATGATGATGATGATGATGATGATGATG	300
QY	301	TATTAGCAAGACAGCTTTCTTGTGTACTTTTGTGTTTCAAAACCCCATGGTAACTGCCT	360
DB	301	TATTAGCAAGACAGCTTTCTTGTGTACTTTTGTGTTTCAAAACCCCATGGTAACTGCCT	360
QY	361	TAACTTTTGTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	420
DB	361	TAACTTTTGTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	420
QY	421	GTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	480
DB	421	GTTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	480
QY	481	AGGAGAGAGTGGCAAGAGAAACTGGAGAGGACCTGTCCAGGGGGGCTGAAGACA	540
DB	481	AGGAGAGAGTGGCAAGAGAAACTGGAGAGGACCTGTCCAGGGGGGCTGAAGACA	540
QY	541	CCGGACAGTGCACAAAGAGAGCCCAAGAGGATAGTGGTCAACCGGACCTGCCCAT	600
DB	541	CCGGACAGTGCACAAAGAGAGCCCAAGAGGATAGTGGTCAACCGGACCTGCCCAT	600
QY	601	GTGACTGTTTCTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCTT	660
DB	601	GTGACTGTTTCTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCTT	660
QY	661	AGCAGTGGGTATGAGTGTGACGGGGCTGGTGGCTTCTTCAGCCATTACAAGAGG	720
DB	661	AGCAGTGGGTATGAGTGTGACGGGGCTGGTGGCTTCTTCAGCCATTACAAGAGG	720
QY	721	GCCCCCACCACCCCGGACGCTGGAGGCTCTGTGTCTCTTAAGCCTCCTTA	780
DB	721	GCCCCCACCACCCCGGACGCTGGAGGCTCTGTGTCTCTTAAGCCTCCTTA	780
QY	781	CTCTCTGGGCTCATGACTATCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG	840
DB	781	CTCTCTGGGCTCATGACTATCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG	840

DB	781	CTCTCTGGGCTCATGACTATCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG	840
QY	841	GACTGTAGTCTGATTTTACTCTGTGAACACTTATTTAAGACATTTCTTTTATTG	900
DB	841	GACTGTAGTCTGATTTTACTCTGTGAACACTTATTTAAGACATTTCTTTTATTG	900
QY	901	CGGCTCTGTGACCCCTAGCGCTTGACCCGCTCTCTGTGTACACTTTCAAGCAACAC	960
DB	901	CGGCTCTGTGACCCCTAGCGCTTGACCCGCTCTCTGTGTACACTTTCAAGCAACAC	960
QY	961	TTTTTTCAGACTAAAGGCCAAACAAAGCTAATCGTCTCATAGTGTCATGCTTACTCTC	1020
DB	961	TTTTTTCAGACTAAAGGCCAAACAAAGCTAATCGTCTCATAGTGTCATGCTTACTCTC	1020
QY	1021	CTACCTTGGCCCCATATGTGTGGCGGCTGTGTCTGCATGCTTGTGTGTCTCCGAGCC	1080
DB	1021	CTACCTTGGCCCCATATGTGTGGCGGCTGTGTCTGCATGCTTGTGTGTCTCCGAGCC	1080
QY	1081	TCTGGAGAGGGGGCAGTGAATGTGGAGCAGGAGCTGGGAGCTTTAACTCCTGAGTAAGCA	1140
DB	1081	TCTGGAGAGGGGGCAGTGAATGTGGAGCAGGAGCTGGGAGCTTTAACTCCTGAGTAAGCA	1140
QY	1141	GCCGCTGGAGGCCATCTGCGCTGTTCGCTTTCAGGACGACTTCTCTAGTTCGTATATG	1200
DB	1141	GCCGCTGGAGGCCATCTGCGCTGTTCGCTTTCAGGACGACTTCTCTAGTTCGTATATG	1200
QY	1201	GAGGGGGCAGGCTTACCACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1260
DB	1201	GAGGGGGCAGGCTTACCACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1260
QY	1261	GTTTGGCTTAGCAGCGCATCAATCACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1320
DB	1261	GTTTGGCTTAGCAGCGCATCAATCACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1320
QY	1321	ACAGCCCATGAACCTACGATCCCGTGTGTCTGCGGTGGCGGAAGCGGAGCGGTAAG	1380
DB	1321	ACAGCCCATGAACCTACGATCCCGTGTGTCTGCGGTGGCGGAAGCGGAGCGGTAAG	1380
QY	1381	GAGGTACACAGCTGTCTTTCGAGAGGGGGTAGGGGGCTCCATGAATGGAAGCGGCGGC	1440
DB	1381	GAGGTACACAGCTGTCTTTCGAGAGGGGGTAGGGGGCTCCATGAATGGAAGCGGCGGC	1440
QY	1441	GGCGGAGCAGCTGAGCTGGATTCGCGGGCCGGGCGAGGGGCTGCCAGGCGCCGACCC	1500
DB	1441	GGCGGAGCAGCTGAGCTGGATTCGCGGGCCGGGCGAGGGGCTGCCAGGCGCCGACCC	1500
QY	1501	GTGTATGGGGCGGCTTCTGTGATCCTTAAGAGCAAGGACCGACCGGCGGCTCAGATGCC	1560
DB	1501	GTGTATGGGGCGGCTTCTGTGATCCTTAAGAGCAAGGACCGACCGGCGGCTCAGATGCC	1560
QY	1561	GGGGCGGGCCGGGAGGCTCGGGCGGAGGCGGTGGGGCTGGAGCGGCTCAGATGCC	1620
DB	1561	GGGGCGGGCCGGGAGGCTCGGGCGGAGGCGGTGGGGCTGGAGCGGCTCAGATGCC	1620
QY	1621	CGGAAGGAAGCAAGCCATCCACAGCCCGTGAATGGAGGGGGTCAAAAGGTCAAGGGGTC	1680
DB	1621	CGGAAGGAAGCAAGCCATCCACAGCCCGTGAATGGAGGGGGTCAAAAGGTCAAGGGGTC	1680
QY	1681	AGGGGCTTGAAGCCGGAGGAAACAGGGGTGGGTCAGTAGAGTGGGCTCAGGTCAAGGTC	1740
DB	1681	AGGGGCTTGAAGCCGGAGGAAACAGGGGTGGGTCAGTAGAGTGGGCTCAGGTCAAGGTC	1740
QY	1741	GAGGGGACTCTCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCTCTCTTACCAGAGTA	1800
DB	1741	GAGGGGACTCTCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCTCTCTTACCAGAGTA	1800
QY	1801	TTAAAGGAACCTGAGGGTCAAGAGTACGGGAAGTGAAGTTCACAAAGCTTGGCTCTTGT	1860
DB	1801	TTAAAGGAACCTGAGGGTCAAGAGTACGGGAAGTGAAGTTCACAAAGCTTGGCTCTTGT	1860
QY	1861	GTTTCGATTTATGGGTACTGTGTGGAGGAGATTCCACAAAGCACCCTCCCTCTTACTGT	1920
DB	1861	GTTTCGATTTATGGGTACTGTGTGGAGGAGATTCCACAAAGCACCCTCCCTCTTACTGT	1920

QY	1921	GATCTGAGATTCCTTCGAGAAAGTTCCTGGGCGCTTGGTGGAGAGCGATTGCTA	1980
Db	1921	GATCTGAGATTCCTTCGAGAAAGTTCCTGGGCGCTTGGTGGAGAGCGATTGCTA	1980
QY	1981	CTAGCCCAAGATCTGAGATTTTGAAGAGAAAGAGAGAGAGTGAAGTGACGGCCACAAC	2040
Db	1981	CTAGCCCAAGATCTGAGATTTTGAAGAGAAAGAGAGAGAGTGAAGTGACGGCCACAAC	2040
QY	2041	GACAGCTCANTGGCTTTGAGAGAGACTCTGAGAGAGACTCTCAGGGGGCCAGACCTTGA	2100
Db	2041	GACAGCTCANTGGCTTTGAGAGAGACTCTGAGAGAGACTCTCAGGGGGCCAGACCTTGA	2100
QY	2101	CTTCCCTATGGGCTGAGTGAACAGAGTCTGGGGGCGCCGCGCACTAAGTGCGAGAGT	2160
Db	2101	CTTCCCTATGGGCTGAGTGAACAGAGTCTGGGGGCGCCGCGCACTAAGTGCGAGAGT	2160
QY	2161	GAAGTTGAGAAACAGACAGGGGTCCAGGGAGGCGCAGGGGTGAGAGCCAGGGCCAGCC	2220
Db	2161	GAAGTTGAGAAACAGACAGGGGTCCAGGGAGGCGCAGGGGTGAGAGCCAGGGCCAGCC	2220
QY	2221	TGTCAGCTGTGTGGGGGCGGACAGTGAAGGGGCGGTGTGTGGGGCAGAGAGGGCGGGT	2280
Db	2221	TGTCAGCTGTGTGGGGGCGGACAGTGAAGGGGCGGTGTGTGGGGCAGAGAGGGCGGGT	2280
QY	2281	GGGGGGCCCCCGCTGCCCCCA CGGTTCATGTCGGGCTGTGCGCTTTGCTGTCC	2340
Db	2281	GGGGGGCCCCCGCTGCCCCCA CGGTTCATGTCGGGCTGTGCGCTTTGCTGTCC	2340
QY	2341	CACCTACTCAGCACTGAAGCGGCACATGCGAGACACACAGCGGGAGAGCGGTTCGCG	2400
Db	2341	CACCTACTCAGCACTGAAGCGGCACATGCGAGACACACAGCGGGAGAGCGGTTCGCG	2400
QY	2401	TGTGGCGCGTGGCCATPACGNTCAGCCAGTTGCTCACTGACGCGACATACCGGACCC	2460
Db	2401	TGTGGCGCGTGGCCATPACGNTCAGCCAGTTGCTCACTGACGCGACATACCGGACCC	2460
QY	2461	CATCTGGGAGAGCCCTACCGTGTGCCCTCCAGTCCCTTGGCTGACGAGCTTGCGC	2520
Db	2461	CATCTGGGAGAGCCCTACCGTGTGCCCTCCAGTCCCTTGGCTGACGAGCTTGCGC	2520
QY	2521	AACCTGAGGGGATCAGGGCACCCACACAGGGGCTCCCACTCTCTCCCTGCCCAACCTGT	2580
Db	2521	AACCTGAGGGGATCAGGGCACCCACACAGGGGCTCCCACTCTCTCCCTGCCCAACCTGT	2580
QY	2581	GGCTTCGATGCTGCTCAGACCAACCCGGCTCCAGTCCCAAGCAGAGAGAGGGG	2640
Db	2581	GGCTTCGATGCTGCTCAGACCAACCCGGCTCCAGTCCCAAGCAGAGAGAGGGG	2640
QY	2641	ACAATGCCCGAGATCAGAAATGCGCTGATCCTGCCAGACTTGAGTCTTCATGTGCCA	2700
Db	2641	ACAATGCCCGAGATCAGAAATGCGCTGATCCTGCCAGACTTGAGTCTTCATGTGCCA	2700
QY	2701	CCAGGTGGTCCAGTTTCCTGCGCAGACTGTGGGAGCTGGGGGTGAAGGGAGAGCTTG	2760
Db	2701	CCAGGTGGTCCAGTTTCCTGCGCAGACTGTGGGAGCTGGGGGTGAAGGGAGAGCTTG	2760
QY	2761	TGTGAACTCGATCCGAACCACTGCGAGACTACTGTTCCTTTGGACCTGCCGGGCTGT	2820
Db	2761	TGTGAACTCGATCCGAACCACTGCGAGACTACTGTTCCTTTGGACCTGCCGGGCTGT	2820
QY	2821	GGACAGAACTGGAGAGGGTGAAGGCGAGCGTGGAGCTGCCATGTGTGGGCGCTGC	2880
Db	2821	GGACAGAACTGGAGAGGGTGAAGGCGAGCGTGGAGCTGCCATGTGTGGGCGCTGC	2880
QY	2881	ATGCGAGAGAGCTGAGGGGTGCGACTGGGGGACCCAGGGCCCTGGTGAACAAGGC	2940
Db	2881	ATGCGAGAGAGCTGAGGGGTGCGACTGGGGGACCCAGGGCCCTGGTGAACAAGGC	2940
QY	2941	TTTGCTGTAGTTTATGCCCTTTTGGCACTCACTACCCCAACCACTGGCTCGGCAATG	3000
Db	2941	TTTGCTGTAGTTTATGCCCTTTTGGCACTCACTACCCCAACCACTGGCTCGGCAATG	3000

QY	3001	AAAGCTCAGTGGTGAGAAACCTTTCGGCTGTGCCGCTGTCCATACGCTCTGCTCAT	3060
Db	3001	AAAGCTCAGTGGTGAGAAACCTTTCGGCTGTGCCGCTGTCCATACGCTCTGCTCAT	3060
QY	3061	CTGATAAACCCTGAAACCGCACAGCGCGTCCACACAGGAGAAAGCCCTACAGTGCCCC	3120
Db	3061	CTGATAAACCCTGAAACCGCACAGCGCGTCCACACAGGAGAAAGCCCTACAGTGCCCC	3120
QY	3121	CTCTGTCGATGCTGCTGAGCAACCTTGGCAACCTCAAGCGTCATGTCGATCCACTCT	3180
Db	3121	CTCTGTCGATGCTGCTGAGCAACCTTGGCAACCTCAAGCGTCATGTCGATCCACTCT	3180
QY	3181	GGTGACAAACCTTTTCGGTGTAGCTTTCGCACTACAGCTGCAACACAGAGTATGAACCTC	3240
Db	3181	GGTGACAAACCTTTTCGGTGTAGCTTTCGCACTACAGCTGCAACACAGAGTATGAACCTC	3240
QY	3241	AAACGTCATATGCTGGGACACACGGGCGAGAACCTTCGCTGTGCGCACTCGCGCTAT	3300
Db	3241	AAACGTCATATGCTGGGACACACGGGCGAGAACCTTCGCTGTGCGCACTCGCGCTAT	3300
QY	3301	ACCACAGGCCACTGGGACACTACAGCGTCATCAGAGGTGCATGCGCATGTGGAGCA	3360
Db	3301	ACCACAGGCCACTGGGACACTACAGCGTCATCAGAGGTGCATGCGCATGTGGAGCA	3360
QY	3361	GGAGGGCTGCTCTCTGCGCCCTGAGGGCTGGGCGCCCACTCATAGCCACCCCTCTGTT	3420
Db	3361	GGAGGGCTGCTCTCTGCGCCCTGAGGGCTGGGCGCCCACTCATAGCCACCCCTCTGTT	3420
QY	3421	TTGAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTCATTCAGAC	3480
Db	3421	TTGAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTCATTCAGAC	3480
QY	3481	TCACCTTGAATTAAGTCTAGTTCTTTTACCTGGGCTCTAGGAAATAGCCCTATGCTCT	3540
Db	3481	TCACCTTGAATTAAGTCTAGTTCTTTTACCTGGGCTCTAGGAAATAGCCCTATGCTCT	3540
QY	3541	GCATTTTATACAAATGAACATAGAAACCACTTCCCTTCCCTTCTCCCGCTGGTCAAGGCT	3600
Db	3541	GCATTTTATACAAATGAACATAGAAACCACTTCCCTTCCCTTCTCCCGCTGGTCAAGGCT	3600
QY	3601	CCACACAGCTAACTAGGCACTATATGGACCGCTGAATCCCATGCTCAGGGGGCCAT	3660
Db	3601	CCACACAGCTAACTAGGCACTATATGGACCGCTGAATCCCATGCTCAGGGGGCCAT	3660
QY	3661	ATAGACCAAGGGACTTGTCTTAGCTCAGTACCGATGAGCTAAGTATGAGGCCCTTGG	3720
Db	3661	ATAGACCAAGGGACTTGTCTTAGCTCAGTACCGATGAGCTAAGTATGAGGCCCTTGG	3720
QY	3721	ATTACCGCCACTGCTCCAGAGGCTATGGATGAATGTTGGAGCTGCCAGCCTTTT	3780
Db	3721	ATTACCGCCACTGCTCCAGAGGCTATGGATGAATGTTGGAGCTGCCAGCCTTTT	3780
QY	3781	ACTGTTTTAACTTATTTTCAGTGTCTTATAATAAAGGAAACACTAACAGAAAAA	3840
Db	3781	ACTGTTTTAACTTATTTTCAGTGTCTTATAATAAAGGAAACACTAACAGAAAAA	3840
QY	3841	AAAAAAA 3847	
Db	3841	AAAAAAA 3847	

RESULT 2

AAZ35053

ID AAZ35053 standard; cDNA; 2651 BP.

XX

AC AAZ35053;

XX

DT 28-FEB-2000 (first entry)

XX

DE Mouse CNREB-2L (long isoform) cDNA.

XX

KW CNREB-2L; CNRE binding factor; transcription factor; mouse; renin;
negative regulatory element; cardiovascular disease; hypertension;

KW congestive heart failure; diagnosis; therapy; ss.
XX Mus musculus.
OS
XX
FH Key Location/Qualifiers
FT CDS 674..2293
FT /*tag= a
FT polyA_signal 2613..2618
FT /*tag= b
XX
XX
XX WO9955343-A1.
XX
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-US008502.
XX
XX 24-APR-1998; 98US-0082997P.
XX (BGM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Chen YE, Horiuchi M, Dzau VJ, Tamura K;
XX
XX WPI; 2000-062002/05.
XX P-PSDB; AAY32377.
XX
XX A new regulatory element binding factor used to treat an adverse
XX cardiovascular condition.
XX
XX
XX Claim 1; Page 85-86; 57pp; English.
XX
XX This is the nucleotide sequence of a cDNA clone encoding mouse negative
XX regulatory element binding transcription factor 2 (CNREB-2) long isoform
XX CNREB-2L. The clone was obtained using yeast one-hybrid cloning and mouse
XX kidney cDNA library screening. CNREB-2 binds to the 5' flanking region of
XX the renin gene and modulates expression of the gene. 2 Other isoforms
XX (see AAY32375 and AAY32376) of CNREB-2 have been identified, the short
XX isoform showing strongest binding activity to the CNRE of the renin
XX promoter. CNREB-2 shares no significant homology to any other known gene
XX or protein. It is a novel zinc finger transcription factor. Expression
XX vectors, isolated nucleic acid molecules and CNREB-2 polypeptides are
XX claimed. The invention also relates to CNREB-1 polynucleotides (see
XX AA235040-43), polypeptides (see AAY32374) and inhibitors, and their use
XX in claimed methods for decreasing or increasing renin expression, and for
XX determining the level of CNREB-1 expression, an increase in expression
XX being indicative of an increased susceptibility to development of a renin
XX -angiotensin system mediated disorder
XX
XX Sequence 2651 BP; 548 A; 744 C; 851 G; 508 T; 0 U; 0 Other;
XX
XX
XX Query Match 68.3%; Score 2628.4; DB 3; Length 2651;
XX Best Local Similarity 99.7%; Pred. No. 0;
XX Matches 2643; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
XX
XX 1198 ATCGAGGGGCGAGGTACACACTTCTGGTTGGCTGGTTACGGTCCAGCTTCG 1257
XX 1 ATGGAGGGGCGAGGTACACACTTCTGGTTGGCTGGTTACGGTCCAGCTTCG 60
XX
XX 1258 ATTGGTTGGCTAGCAGCGCATCACAAATCACTTCCTGGTCACTGCTCGACTCGCTCCGGT 1317
XX 61 ATTGGTTGGCTAGCAGCGCATCACAAATCACTTCCTGGTCACTGCTCGACTCGCTCCGGT 120
XX
XX 1318 TGTAAGCCCATGAAGTACGATCCGTTGCTCTCGGGTGGCGGAGCGGAGCGGTT 1377
XX 121 TGTACAGCCCATGAAGTACGATCCGTTGCTCTCGGGTGGCGGAGCGGAGCGGTT 180
XX
XX 1378 ACGAGGTACCAAGTGGTCTTCGAGGGGGGTAGGGGGTCCATGAATGGAAGCGGCGGC 1437
XX 181 ACGAGGTACCAAGTGGTCTTCGAGGGGGGTAGGGGGTCCATGAATGGAAGCGGCGGC 240
XX
XX 1438 GCGGGGCGGAGGACCTGAGCTGGATTCGCGGGCGCGGG- CAGGGGCTGCCAGGGGCGCG 1496
XX 241 GCGGGGCGGAGGACCTGAGCTGGGCTCGGGGCGCGGGCCAGGGGCTGCCAGGGGCGCG 300

QY 1497 CACCGGTATATGGGCGGGTTCGTGATCTTAAGACAAAGGACCGACGGAGGGCGGAACCT 1556
DB 301 CACCGGTATATGGGCGGGTTCGTGATCTTAAGACAAAGGACCGACGGAGGGCGGAACCT 360
QY 1557 GGGAGGGGCGGGCGGGAGGCTTCGGGCGGAGCGCTCGGGCTGGAGCGGCTACAGCA 1616
DB 361 GGGAGGGGCGGGCGGGAGGCTTCGGGCGGAGCGGCTCGGGCTGGAGCGGCTACAGCA 420
QY 1617 TGCCCGGAAGGAAGCAAGCCATCCACAGCCCGTGAATTCGAGAGGGGTCAAAGTCAAG 1676
DB 421 TGCCCGGAAGGAAGCAAGCCATCCACAGCCCGTGAATTCGAGAGGGGTCAAAGTCAAG 480
QY 1677 GGTTCAGGGGCGGCTTCAGCGGGGAGGACAGGGGTGGGTTCAGTAGAGTGGGTTCAGGTTCAG 1736
DB 481 GGTTCAGGGGCGGCTTCAGCGGGGAGGACAGGGGTGGGTTCAGTAGAGTGGGTTCAGGTTCAG 540
QY 1737 GGTTCAGGGGCGGCTTCAGGGGTTCAGGGGCGGATCTTCGGATCTTCGCTCCTTACCG 1796
DB 541 GGTTCAGGGGCGGCTTCAGGGGTTCAGGGGCGGATCTTCGGATCTTCGCTCCTTACCG 600
QY 1797 AGTATTAAAGGAACCTGAGGTTCAGGATCGGGAAGTGCAGTTTCACAACAGCTGGGTC 1856
DB 601 AGTATTAAAGGAACCTGAGGTTCAGGATCGGGAAGTGCAGTTTCACAACAGCTGGGTC 660
QY 1857 CTTGGTTCCGATTATGGGTACTGCTTCGAGGGGAGATTCCACAAGCACCTCCCTCTTT 1916
DB 661 CTTGGTTCCGATTATGGGTACTGCTTCGAGGGGAGATTCCACAAGCACCTCCCTCTTT 720
QY 1917 AGTGGATACCTAAGATTCCTTCGACGAGGTCTCGGGGCGCTTCGGTTCGGAGAGCGATT 1976
DB 721 AGTGGATACCTAAGATTCCTTCGACGAGGTCTCGGGGCGCTTCGGTTCGGAGAGCGATT 780
QY 1977 GCTACTAGGCCAAGATCTCGAGTTTCAAGAGGAAGAGGATGAAGTGAAGGCCA 2036
DB 781 GCTACTAGGCCAAGATCTCGAGTTTCAAGAGGAAGAGGATGAAGTGAAGGCCA 840
QY 2037 CAACGACCAAGCTCATGGGTTTGAGAGACCTCTGAAGGAGACTCTCAGGGGGCGGAGCC 2096
DB 841 CAACGACCAAGCTCATGGGTTTGAGAGACCTCTGAAGGAGACTCTCAGGGGGCGGAGCC 900
QY 2097 TGACCTTCCTATGGGCTGAGTTCAGACCAAGTCTGGGGCGGCGGCTAAGTGCAG 2156
DB 901 TGACCTTCCTATGGGCTGAGTTCAGACCAAGTCTGGGGCGGCGGCTAAGTGCAG 960
QY 2157 GAGTGAAGTTGAGGAACCAAGCCAGGGGTTCAGGGGAGGCCAGGGGTGAGAGCCAGGCC 2216
DB 961 GAGTGAAGTTGAGGAACCAAGCCAGGGGTTCAGGGGAGGCCAGGGGTGAGAGCCAGGCC 1020
QY 2217 AGCTGTACGCTGTGGGGGCGGACAGGTGAGGGGCGGCTGTGTGGGCGAGGGGCG 2276
DB 1021 AGCTGTACGCTGTGGGGGCGGACAGGTGAGGGGCGGCTGTGTGGGCGAGGGGCG 1080
QY 2277 GGGTGGGGGCGGCGGCTGCCCCCAGGTTACTGTACTCATGCGGCTGTGGCTTCGT 2336
DB 1081 GGGTGGGGGCGGCGGCTGCCCCCAGGTTACTGTACTCATGCGGCTGTGGCTTCGT 1140
QY 2337 GTCCCACTACTCAGCGCACTGAGCGGCACATGACACACAGCGGGAGAGCGGTT 2396
DB 1141 GTCCCACTACTCAGCGCACTGAGCGGCACATGACACACAGCGGGAGAGCGGTT 1200
QY 2397 CCGCTGTGGCGGCTGCGCATACGCNTCAGCCCAAGTTCTGTCACCTCAGCGCATATACCG 2456
DB 1201 CCGCTGTGGCGGCTGCGCATACGCCTCAGCCCAAGTTCTGTCACCTCAGCGCATATACCG 1260
QY 2457 CACCCATACTGGCGAGAACCCCTTACCGTTGTCCCACTGCGGCTGTGGCTGAGAGCGCT 2516
DB 1261 CACCCATACTGGCGAGAACCCCTTACCGTTGTCCCACTGCGGCTGTGGCTGAGAGCGCT 1320
QY 2517 GGGCACTGAGCGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTCCCTGCCCAAC 2576
DB 1321 GGGCACTGAGCGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTCCCTGCCCAAC 1380
QY 2577 CTGTGGCTTTTCGATGTGTGCTCCACGACCAACCCGCGCTCCCACTCCCAAGAGCAGGA 2636

1381 CTGTGGCTTTTCATGCTGTGCTCCACGACCAACCCGGGCTCCAGTCCACAGAGCAGGA 1440
2637 GGGGCAATGCCGCCACGATCAGAAAATGCGTGAATCCTCGCAGACTTGAGTTCATGT 2696
1441 GGGGCAATGCCGCCACGATCAGAAATGCGTGAATCCTCGCAGACTTGAGTTCATGT 1500
2697 GCCACAGTGTGTCAGTTTCTGTCAGACTGTGTGGGAGCTGCGGGTGAAGGGAGAG 2756
1501 GCCACAGTGTGTCAGTTTCTGTCAGACTGTGTGGGAGCTGCGGGTGAAGGGAGAG 1560
2757 CTGTGTGGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 2816
1561 CTGTGTGGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 1620
2817 CTGTGACAGGAATCGGAGAGGGTGAAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 2876
1621 CTGTGACAGGAATCGGAGAGGGTGAAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 1680
2877 CTGCAATCGGAGAGAGGCTGGAAGGGTTCACCTGGGGAGCCCGAGGCGCTGTGACAA 2936
1681 CTGCAATCGGAGAGAGGCTGGAAGGGTTCACCTGGGGAGCCCGAGGCGCTGTGACAA 1740
2937 AGGCTTTGCTGTAGTTATGCCCTTTGCCACTCACTACCCCAACCACTGGCTCGGCA 2996
1741 AGGCTTTGCTGTAGTTATGCCCTTTGCCACTCACTACCCCAACCACTGGCTCGGCA 1800
2997 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTCGCGCTTCCATAGCCTCTGC 3056
1801 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTCGCGCTTCCATAGCCTCTGC 1860
3057 TCATCTGGATTAACCTGAAACGGCAGCAGCGCTGCCACAGAGAGAAAGCCCTACAAGTG 3116
1861 TCATCTGGATTAACCTGAAACGGCAGCAGCGCTGCCACAGAGAGAAAGCCCTACAAGTG 1920
3117 CCCCCTCTGCTGATGCTGTGGAACTTGGCAACCTTGGCAACCTTGGCAACCTTGGCAAC 3176
1921 CCCCCTCTGCTGATGCTGTGGAACTTGGCAACCTTGGCAACCTTGGCAACCTTGGCAAC 1980
3177 CTCTGTGTGACAAACCTTTTGGGTGAGCTTGGCAACCTTGGCAACCTTGGCAACCTTGGCA 3236
1981 CTCTGTGTGACAAACCTTTTGGGTGAGCTTGGCAACCTTGGCAACCTTGGCAACCTTGGCA 2040
3237 CCTCAACGCTCATATGCTGCGACACACGGGCGAGAAAGCCCTTCGCTGTGCGCACTTGCGC 3296
2041 CCTCAACGCTCATATGCTGCGACACACGGGCGAGAAAGCCCTTCGCTGTGCGCACTTGCGC 2100
3297 CTATACACAGGCGCATGCGGACAACTACAGCGCTCATCAGAGGTGCTAGGCGCATGCTGG 3356
2101 CTATACACAGGCGCATGCGGACAACTACAGCGCTCATCAGAGGTGCTAGGCGCATGCTGG 2160
3357 AGCAGAGGGGCTGTCTCTCTGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGG 3416
2161 AGCAGAGGGGCTGTCTCTCTGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGG 2220
3417 TGTTTTGAGCACTCGGGGTCACAGCGCTGGGTGCTACTGTAGCAGGGCTCTTCATTC 3476
2221 TGTTTTGAGCACTCGGGGTCACAGCGCTGGGTGCTACTGTAGCAGGGCTCTTCATTC 2280
3477 AGACTCACTTTGAATTAACCTAGTGTCTTTTACCTGGGCTCTAGGAATTAGCCTATGCC 3536
2281 AGACTCACTTTGAATTAACCTAGTGTCTTTTACCTGGGCTCTAGGAATTAGCCTATGCC 2340
3537 TCCTGCAATTTATACAAATGAATAGAAACCACTTCTCCCTTCTCCCGCTGTGTCAGG 3596
2341 TCCTGCAATTTATACAAATGAATAGAAACCACTTCTCCCTTCTCCCGCTGTGTCAGG 2400
3597 GGCTTCCACAGACTTAACCTAGGCACTATATGACAGGCTGAAATCCCTGTTGTCAGGGG 3656
2401 GGCTTCCACAGACTTAACCTAGGCACTATATGACAGGCTGAAATCCCTGTTGTCAGGGG 2460
3657 CCATATAGACAGGGGACTTGTCTTACCTCAGCTCAGTACAGATAGCTAAGTATAGGGCC 3716

2461 CCATATAGACAGGGGACTTGTCTTAGCTCACGTACCAGATAGCTAAGTATAGGGCC 2520
3717 TTGGATTACCGCACTGTCTCCAGAGGCTATGGATGAACCTGTTGGAGCTCCCGAGCC 3776
2521 TTGGATTACCGCACTGTCTCCAGAGGCTATGGATGAACCTGTTGGAGCTCCCGAGCC 2580
3777 TTTTACTGTTTTTAACCTATTTCAGTGTCTTTATATAAGGAAACACTTAACAGAAAAAAA 3836
2581 TTTTACTGTTTTTAACCTATTTCAGTGTCTTTATATAAGGAAACACTTAACAGAAAAAAA 2640
3837 AAAAAA 3847
2641 AAAAAA 2651

RESULT 3
AAZ35052
ID AAZ35052 standard; cDNA; 2336 BP.
XX AAZ35052;
AC
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse CNREB-2S (short isoform) cDNA.
KW CNREB-2S; CNRE binding factor; transcription factor; mouse; renin;
KW negative regulatory element; cardiovascular disease; hypertension;
KW congestive heart failure; diagnosis; therapy; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 674..1978
FT /*tag= a
FT polyA_signal 2298..2303
FT /*tag= b
XX
PN WO955343-A1.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US008502.
XX
PR 24-APR-1998; 98US-0082997P.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Chen YE, Horiuchi M, Dzau VJ, Tamura K;
XX
XX WPI: 2000-062002/05.
DR P-PSDB; AAY32376.
XX
XX A new regulatory element binding factor used to treat an adverse
cardiovascular condition.
XX
XX Claim 1; Page 83-84; 57pp; English.
XX
XX This is the nucleotide sequence of a cDNA clone encoding mouse negative
regulatory element binding transcription factor 2 (CNREB-2) short isoform
CNREB-2S. The clone was obtained using yeast one-hybrid cloning and mouse
kidney cDNA library screening. CNREB-2 binds to the 5' flanking region of
the renin gene and modulates expression of the gene. 2 Other isoforms
(see AAY32375 and AAY32377) of CNREB-2 have been identified, the short
isoform showing strongest binding activity to the CNRE of the renin
promoter. CNREB-2 shares no significant homology to any other known gene
or protein. It is a novel zinc finger transcription factor. Expression
vectors, isolated nucleic acid molecules and CNREB-2 polypeptides are
claimed. The invention also relates to CNREB-1 polynucleotides (see
AAZ35040-43), polypeptides (see AAY32374) and inhibitors, and their use
in claimed methods for decreasing or increasing renin expression, and for
determining the level of CNREB-1 expression, an increase in expression
being indicative of an increased susceptibility to development of a renin
-angiotensin system mediated disorder

XX SQ Sequence 2336 BP; 494 A; 661 C; 737 G; 444 T; 0 U; 0 Other;

Query Match 51.7%; Score 1988.4; DB 3; Length 2336;
 Best Local Similarity 87.8%; Pred. No. 0;
 Matches 2328; Conservative 0; Mismatches 7; Indels 316; Gaps 2;

QY 1198 ATGAGGGGGGAGAGGTACACCTTCCTGTTGGCTGGTTTACCGTCCAGCTTCG 1257
 DB 1 ATGAGGGGGGAGAGGTACACCTTCCTGTTGGCTGGTTTACCGTCCAGCTTCG 60

QY 1258 ATTGTTGGCTAGAGGCGCATCACATCACTTCCTGTCATGCTGGAGCTGCGCTCCGGT 1317
 DB 61 ATTGTTGGCTAGAGGCGCATCACATCACTTCCTGTCATGCTGGAGCTGCGCTCCGGT 120

QY 1318 TGTACAGCCCATGAATACGATCCCGTGTGCTCTGCGGTGGCGGAAGCGGAGCGGGT 1377
 DB 121 TGTACAGCCCATGAATACGATCCCGTGTGCTCTGCGGTGGCGGAAGCGGAGCGGGT 180

QY 1378 ACGAGGTACACAGCTGGTCTTCGAGGGGGGTAGGGGCTCCATGATGGAAGCGGGCGG 1437
 DB 181 ACGAGGTACACAGCTGGTCTTCGAGGGGGGTAGGGGCTCCATGATGGAAGCGGGCGG 240

QY 1438 GCGCGGGGAGCGACTCAGCTGGATCCGGGCGCGGG - CAGGGGCTGCCAGGGCGCG 1496
 DB 241 GCGCGGGGAGCGACTCAGCTGGCTCCGGGCGCGGGCCAGGGGCTGCCAGGGGCGG 300

QY 1497 CACCGTGTATGGGGCGGTTCGTGGATCCTAAGAGCAAGAGCCGAGCGGCGCAACT 1556
 DB 301 CACCGTGTATGGGGCGGTTCGTGGATCCTAAGAGCAAGAGCCGAGCGGCGCAACT 360

QY 1557 GGGAGGGCGGGCGGGAGCGCTCGGCGCGAGGCGGCTGGGCTGGAGCGGCTCAGCA 1616
 DB 361 GGGAGGGCGGGCGGGAGCGCTCGGCGCGAGGCGGCTGGGCTGGAGCGGCTCAGCA 420

QY 1617 TGCCCCGGAAGCAAGCAAGCCATCCACAGCCCGTGAATGCGAGGGGTCAAGAGTCAGG 1676
 DB 421 TGCCCCGGAAGCAAGCAAGCCATCCACAGCCCGTGAATGCGAGGGGTCAAGAGTCAGG 480

QY 1677 GGTACAGGGGCTTAGCGGGAGGAGCAAGGGTGGGGTCAGTAGAGTGGGCTCAGGTCAG 1736
 DB 481 GGTACAGGGGCTTAGCGGGAGGAGCAAGGGTGGGGTCAGTAGAGTGGGCTCAGGTCAG 540

QY 1737 GGTACAGGGGACTCCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCGCTCCTTACCAG 1796
 DB 541 GGTACAGGGGACTCCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCGCTCCTTACCAG 600

QY 1797 AGTATTAAGGAACCTGAGGGTCTAGAGTACGGAAATGCGAGTTCAACAGCTGGCTC 1856
 DB 601 AGTATTAAGGAACCTGAGGGTCTAGAGTACGGAAATGCGAGTTCAACAGCTGGCTC 660

QY 1857 CTTGGTTCGGATTATGGGTACTGCTTGGAGGAGATTCCCAAGCACCTCCCTCTTT 1916
 DB 661 CTTGGTTCGGATTATGGGTACTGCTTGGAGGGAGATTCCCAAGCACCTCCCTCTTT 720

QY 1917 AGTGATACGAGATTCCTTCGAGAGGTCCTGGGCGCTGGTGGTGGAGAGCGATTT 1976
 DB 721 AGTGATACGAGATTCCTTCGAGAGGTCCTGGGCGCTGGTGGTGGAGAGCGATTT 780

QY 1977 GCTATAGGCCAGATCTGGAGTTTGAAGAGAGAGGAGGATGAGTGAAGGCCA 2036
 DB 781 GCTATAGGCCAGATCTGGAGTTTGAAGAGAGAGGAGGATGAGTGAAGGCCA 840

QY 2037 CAACGACCAAGCTCATGGGCTTTGAGAGAGCTTGAAGAGAGCTCTCAGGGGGCCAGACC 2096
 DB 841 CAACGACCAAGCTCATGGGCTTTGAGAGAGCTTGAAGAGAGCTCTCAGGGGGCCAGACC 900

QY 2097 TGGACTTCCTATGGGCTAGTGAAGAGCTCTGGGGCGCGCGGCTCAAGTGGCGGA 2156
 DB 901 TGGACTTCCTATGGGCTAGTGAAGAGCTCTGGGGCGCGCGGCTCAAGTGGCGGA 960

QY 2157 GAGTGAAGTTGAGGAACCAAGCGGGGTCAGGGAGGCGGAGGGGTGAGAGGCCAGGGCCC 2216

DB 961 GAGTGAAGTTGAGGAACCAAGCCAGCGGTCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC 1020

QY 2217 AGCCTGTAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTGTGTGGGGAGGAGGCGG 2276

DB 1021 AGCCTGTAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTGTGTGGGGAGGAGGCGG 1080

QY 2277 GGGTGGGGGGCCCCCGCTGCGCCCAACCGGTACTGCTACTGCTGCGGCTGTGGGCTTTGCT 2336

DB 1081 GGGTGGGGGGCCCCCGCTGCGCCCAACCGGTACTGCTACTGCTGCGGCTGTGGGCTTTGCT 1140

QY 2337 GTCCCACTACTCGAGCCACCTGAAGCGGCACATGAGAGCACACACAGCGGGGAGAACCGCTT 2396

DB 1141 GTCCCACTACTCGAGCCACCTGAAGCGGCACATGAGAGCACACACAGCGGGGAGAACCGCTT 1200

QY 2397 CCGCTGTGCGCTGCCCATACGNTCAGCCAGTTCGTTCAACCTGAGCGGACATACCGG 2456

DB 1201 CCGCTGTGCGCTGCCCATACGNTCAGCCAGTTCGTTCAACCTGAGCGGACATACCGG 1260

QY 2457 CACCCATACTCGGAGAGCCCTACGTTGTCGCCCACTGCCCTTTGCTGTCAGCAGCT 2516

DB 1261 CACCCATACTCGGAGAGCCCTACGTTGTCGCCCACTGCCCTTTGCTGTCAGCAGCT 1320

QY 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTGCCAAC 2576

DB 1321 GGGCAACCTGAGCGGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTGCCAAC 1380

QY 2577 CTGTGCTTTCGATGCTGTGCTCCAGCACCCAGCCAGCTCCAGTCCACAGAGCAGGA 2636

DB 1381 CTGTGCTTTCGATGCTGTGCTCCAGCACCCAGCCAGCTCCAGTCCACAGAGCAGGA 1440

QY 2637 GGGGCAATGCCCCCAGCAGTCAAGAAATCGCTGATCTCTGCCAGACTTGGTCTTCATGT 2696

DB 1441 GGGGCAATGCCCCCAGCAGTCAAGAAATCGCTGATCTCTGCCAG - - - - - 1484

QY 2697 GCCACAGGTGGTCCAGTTTCTGCCAGACTGTGGGAGCTGTGGGGTGAAGGGAGAG 2756

DB 1485 - - - - - 1484

QY 2757 CTTGTGTGGAATCGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCGGGG 2816

DB 1485 - - - - - 1484

QY 2817 CTGTGACAGGAACCTGGAGGAGGGTGGAGGAGAGGCTGGGAGCTGGCCATGTGTGGGGG 2876

DB 1485 - - - - - 1484

QY 2877 CTGCATCGAGGAGAGGCTGGAGGGGTGCCACTGGGGGACCCAGGGGCCCTGGTGACAA 2936

DB 1485 - - - - - 1484

QY 2937 AGGCTTTGCTGTAGTTTATGCCCCCTTGGCCACTCACTACCCCAACCACTGGCTCGGCA 2996

DB 1485 - - - - - 1485

QY 2997 CATGAGACTCACAGTGTGGAACCCCTTCCGCTGTGCGCGCTGCCATACGCTCTGC 3056

DB 1486 CATGAGACTCACAGTGTGGAACCCCTTCCGCTGTGCGCGCTGCCATACGCTCTGC 1545

QY 3057 TCATCTGGATAACCTGAAACGGGCACAGCGGCTCCACACAGGAGAAAGCCCTTACAAGTG 3116

DB 1546 TCATCTGGATAACCTGAAACGGGCACAGCGGCTCCACACAGGAGAAAGCCCTTACAAGTG 1605

QY 3117 CCGCTCTGTCCGTATGCTGTGGCAACCTGGCCACCTCAAGCGTCACTGGTGCATCCA 3176

DB 1606 CCGCTCTGTCCGTATGCTGTGGCAACCTGGCCACCTCAAGCGTCACTGGTGCATCCA 1665

QY 3177 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTTCAACTACAGCTGCAACAGAGTATGAA 3236

DB 1666 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTTCAACTACAGCTGCAACAGAGTATGAA 1725

QY 3237 CCTCAACAGTCACTATGCTGCAACACCGGGGAGAGGCCCTTCCGCTGTGCCACCTGCCG 3296

DB 1726 CCTCAACAGTCACTATGCTGCAACACCGGGGAGAGGCCCTTCCGCTGTGCCACCTGCCG 1785

QY 3297 CTATACCAAGCCACTGGGCAACTACAAAGCGTCATCAGAAAGTGCATGCGCATGTTGG 3356
DB 1786 CTATACCAAGCCACTGGGCAACTACAAAGCGTCATCAGAAAGTGCATGCGCATGTTGG 1845
QY 3357 ACAGAGAGGCGCTGCTCTCTGCCCCGAGGCTGGGCCCCACCTCATAGCCACCCCTC 3416
DB 1846 AGCAGAGGCGCTGCTCTCTGCCCCGAGGCTGGGCCCCACCTCATAGCCACCCCTC 1905
QY 3417 TGTGTTTACACTCGGGGTCACAGAGCCCTGGGTGCTACTGTAGCAGAGGCTCTTCATTC 3476
DB 1906 TGTGTTTACACTCGGGGTCACAGAGCCCTGGGTGCTACTGTAGCAGAGGCTCTTCATTC 1965
QY 3477 AGACTCACCCTGAACCTAGTCTTTTACTGCGGCTCTAGAAATTAGCCCTATGCC 3536
DB 1966 AGACTCACCCTGAACCTAGTCTTTTACTGCGGCTCTAGAAATTAGCCCTATGCC 2025
QY 3537 TCTGCAATTTTACAAATGAACCTAGAAACCACTTTCCTTTCTCCCGCTGGTCAAG 3596
DB 2026 TCTGCAATTTTACAAATGAACCTAGAAACCACTTTCCTTTCTCCCGCTGGTCAAG 2085
QY 3597 GGCTCCACACAGACTAAGCTAGGCACTATATGACACAGCTGAATCCCATGTCAGGGG 3656
DB 2086 GGCTCCACACAGACTAAGCTAGGCACTATATGACACAGCTGAATCCCATGTCAGGGG 2145
QY 3657 CCATATAGACCGGAGCTTGTCTAGCTCAGTACACAGCTGAATGATTAGGGCC 3716
DB 2146 CCATATAGACCGGAGCTTGTCTAGCTCAGTACACAGCTGAATGATTAGGGCC 2205
QY 3717 TTGGATTACCGGCACTGCTCCAGAGGCTATGGATGAATGTTGGGAGCTGCCAGGC 3776
DB 2206 TTGGATTACCGGCACTGCTCCAGAGGCTATGGATGAATGTTGGGAGCTGCCAGGC 2265
QY 3777 TTTTACTGTTTAACTTATTCTAGTCTTATAATTAAGGAACACTAACAGAAAAAAA 3836
DB 2266 TTTTACTGTTTAACTTATTCTAGTCTTATAATTAAGGAACACTAACAGAAAAAAA 2325
QY 3837 AAAAAAAAAA 3847
DB 2326 AAAAAAAAAA 2336

RESULT 4

ACC83475
ID ACC83475 standard; cDNA; 1696 BP.
AC ACC83475;
XX 08-SEP-2003 (first entry)
DT Mouse chondrocyte-derived zinc finger protein Czf-1 partial cDNA.
DE Mouse; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritic;
KW antirheumatic; gene; ss.
XX Mus sp.

XX Key Location/Qualifiers
FH 3..1355
CDS /*tag= a
FT /partial
FT /product= "Mouse Czf-1"
FT /note= "No start codon"

XX WO2003044159-A2.

XX 30-MAY-2003.

XX 20-NOV-2002; 2002WO-IL000925.

XX 20-NOV-2001; 2001US-0331626P.

XX (PROC-) PROCHON BIOTECH LTD.

XX Yayan A, Blumenstein S, Harari D;
XX WPI; 2003-457599/43.
DR P-PSDB; ABR42914.
XX New chondrocyte-derived zinc finger polypeptides and encoding
XX polynucleotides, useful for detecting, diagnosing and treating Czf-1
XX protein-related diseases, such as osteoarthritis and rheumatoid
XX arthritis.
XX Example 1; Fig 4; 64pp; English.
XX The present sequence is that of a contig containing most of the murine
XX Czf-1 coding sequence, lacking only the 5' portion which codes for the N-
XX terminus of Czf-1 protein. The contig was obtained by screening mouse
XX expressed sequence tag sequences using rat Czf-1 protein as a sequence
XX probe. Czf-1 is expressed in a variety of tissue types, particularly
XX osteoblasts and chondrocytes. It is expressed in the nuclei of
XX hypertrophic chondrocytes of the growth plate, and exhibits an alternate
XX expression and intracellular localisation pattern in cells expressing
XX high, i.e. constitutively active, levels of fibroblast growth factor
XX receptor (FGFR). It is up-regulated in osteoarthritic cartilage tissue
XX and serves as a marker for osteoarthritis. Czf-1 has been found in rats,
XX mice and humans. Sequence identity for nucleotide and amino acid
XX sequences is above 85%. Czf-1 polynucleotides, polypeptides and
XX antibodies can be used in the characterisation, diagnosis and treatment
XX of FGFR-related and skeletal diseases and disorders, such as
XX osteoarthritis, rheumatoid arthritis, and cartilage related diseases
XX SQ Sequence 1696 BP; 346 A; 529 C; 487 G; 334 T; 0 U; 0 Other;

Query Match 43.5%; Score 1671.8; DB 7; Length 1696;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 2135 GCGGCCCGCACTAAGTTCGGAGAGTGAAGTTGAGGAACCCAGCCAGGGGTCACAGGGAGG 2194
DB 1 GCGGCCCGCACTAAGTTCGGAGAGTGAAGTTGAGGAACCCAGCCAGGGGTCACAGGGAGG 60
QY 2195 CCAGGGGTGAGAGCCAGGCCAGCTGTGCTGCGGCGGCCGACAGAGTTCAGGGGC 2254
DB 61 CCAGGGGTGAGAGCCAGGCCAGCTGTGCTGCGGCGGCCGACAGAGTTCAGGGGC 120
QY 2255 CGTGTGTTGGGGCAGAGAGGCGGGTGGGGGGCCCCCGCTGCCCCACGGTTACTGTTACT 2314
DB 121 CGTGTGTTGGGGCAGAGAGGCGGGTGGGGGGCCCCCGCTGCCCCACGGTTACTGTTACT 180
QY 2315 CATGCCGCTGTGCGCTTTGCTGTCCTACTGAGCCACCTGAGCGGCACATGCGAGA 2374
DB 181 CATGCCGCTGTGCGCTTTGCTGTCCTACTGAGCCACCTGAGCGGCACATGCGAGA 240
QY 2375 CACACAGCGGGGAGAAGCCGTTCCGCTGTGGCGCTGCCCATAGCNTCAGCCAGTTCCG 2434
DB 241 CACACAGCGGGGAGAAGCCGTTCCGCTGTGGCGCTGCCCATAGCNTCAGCCAGTTCCG 300
QY 2435 TCAACCTGACCGACATACCGGACCCATCTGCGGAGAGCCCTAGCGTGTCCCCACT 2494
DB 301 TCAACCTGACCGACATACCGGACCCATCTGCGGAGAGCCCTAGCGTGTCCCCACT 360
QY 2495 GCCCTTTGCTGCGAGAGCTGGGCAACCTGAGGCGGCATCAGCGCACCCACACAGGGC 2554
DB 361 GCCCTTTGCTGCGAGAGCTGGGCAACCTGAGGCGGCATCAGCGCACCCACACAGGGC 420
QY 2555 CTCGACTCTCCCTGCCCAACCTGTGGCTTTCGATGCTGTGCTCCAGGACCAACCGGGC 2614
DB 421 CTCGACTCTCCCTGCCCAACCTGTGGCTTTCGATGCTGTGCTCCAGGACCAACCGGGC 480
QY 2615 CTCGACTCTCCCAAGAGAGGAGGAGCAATGCCCCCGACCATCAGAAAAATGCGTGTATCC 2674
DB 481 CTCGACTCTCCCAAGAGAGGAGGAGGAGCAATGCCCCCGACCATCAGAAAAATGCGTGTATCC 540
QY 2675 TGCCAGACTTGAGTCTTCATGTCACACAGGTGTGCGAGTTTCTCTGCCAGACTGTGGGC 2734

Db 541 TGCCAGACTTGGAGTCTTATGTCACACAGGTGTGCCAGTTTCTCCACAGACTGTGGGC 600
Qy 2735 AGCTCGGGGTGAAGGGAGAGCTTGTGTGAACTGGATCCGAACCACTCCAGAGCTAC 2794
Db 601 AGCTCGGGGTGAAGGGAGAGCTTGTGTGAACTGGATCCGAACCACTCCAGAGCTAC 660
Qy 2795 TGTTCCTTGGACCTCGGGGCTGTGGAACGAACTGGAGAGGGTGGAGGAGCAGGC 2854
Db 661 TGTTCCTTGGACCTCGGGGCTGTGGAACGAACTGGAGAGGGTGGAGGAGCAGGC 720
Qy 2855 TGGAGCTGCCATGTGTGGGCGCTGCATGCGAGGAGGGTGGAGGGGTGCCACTGGGG 2914
Db 721 TGGAGCTGCCATGTGTGGGCGCTGCATGCGAGGAGGGTGGAGGGGTGCCACTGGGG 780
Qy 2915 GA-CCCCAGGCGCTGTGGAACAGGCTTTCCTGTGTATGCCCCCTTGGCACTCAC 2973
Db 781 GACCCCCAGGCGCTGTGGAACAGGCTTTCCTGTGTATGCCCCCTTGGCACTCAC 840
Qy 2974 TACCCCAACCACTGGCTGCGGCATGAAGACTCACAGTGTGAGAAACCTTCCGCTGT 3033
Db 841 TACCCCAACCACTGGCTGCGGCATGAAGACTCACAGTGTGAGAAACCTTCCGCTGT 900
Qy 3034 GCGCGCTGCATACGCTCTGCTCACTGTGATTAACCTGAAACGCGACAGCGGCTCCAC 3093
Db 901 GCGCGCTGCATACGCTCTGCTCACTGTGATTAACCTGAAACGCGACAGCGGCTCCAC 960
Qy 3094 ACAGGAGAAAGCCCTACAGTGGCCCTCTGTCGCTATGCTGTGGCAACCTGGCCCAAC 3153
Db 961 ACAGGAGAAAGCCCTACAGTGGCCCTCTGTCGCTATGCTGTGGCAACCTGGCCCAAC 1020
Qy 3154 CTCAGGCTCATGTGCGCATCCACTCTGTGTGACAAACCTTTCGGTGTAGCCCTTTCGAA 3213
Db 1021 CTCAGGCTCATGTGCGCATCCACTCTGTGTGACAAACCTTTCGGTGTAGCCCTTTCGAA 1080
Qy 3214 TACAGCTGCACACAGATGATGAACCTCAAGCTCATATGCTGCACACACGGGCGAAG 3273
Db 1081 TACAGCTGCACACAGATGATGAACCTCAAGCTCATATGCTGCACACACGGGCGAAG 1140
Qy 3274 CCGTTCCGCTGTGCCACCTCGCCCTATACACAGGCCACTGGGCAACTACAGCGTCTAT 3333
Db 1141 CCGTTCCGCTGTGCCACCTCGCCCTATACACAGGCCACTGGGCAACTACAGCGTCTAT 1200
Qy 3334 CAGAAGTGCATGCGCATGTGTGAGCAGAGGGCTGTCTCTGCCCCCTGAGGGCTGG 3393
Db 1201 CAGAAGTGCATGCGCATGTGTGAGCAGAGGGCTGTCTCTGCCCCCTGAGGGCTGG 1260
Qy 3394 GCCCCACCTCATAGCCACCCCTCTGTTTGTAGCACTCGGGTCCAGCAGCCCTGGGTGCT 3453
Db 1261 GCCCCACCTCATAGCCACCCCTCTGTTTGTAGCACTCGGGTCCAGCAGCCCTGGGTGCT 1320
Qy 3454 ACTGTAGCAGGGCTTTTCATTCAGATCACTTGAACCTAGTGGTCTTTTACCTGGG 3513
Db 1321 ACTGTAGCAGGGCTTTTCATTCAGATCACTTGAACCTAGTGGTCTTTTACCTGGG 1380
Qy 3514 GCTCTAGAAATTAGCCCTATGCCCTCTGCTTATACAAATGAACCTAGAAACCACTTT 3573
Db 1381 GCTCTAGAAATTAGCCCTATGCCCTCTGCTTATACAAATGAACCTAGAAACCACTTT 1440
Qy 3574 CCTTTCTCCCGCTGTGTGAGGGCTCCACACAGACTAAGCTAGGCACTATATGACCA 3633
Db 1441 CCTTTCTCCCGCTGTGTGAGGGCTCCACACAGACTAAGCTAGGCACTATATGACCA 1500
Qy 3634 GCCTGAATCCCATGCTCAGGGGGCCATATACACAGGGGACTTCTTGTAGCTACGTACC 3693
Db 1501 GCCTGAATCCCATGCTCAGGGGGCCATATACACAGGGGACTTCTTGTAGCTACGTACC 1560
Qy 3694 AGATGAGCTAAGTATTAGGCGCTTGGATTACCGCCACTGCTCCAGAGGCTATGATG 3753
Db 1561 AGATGAGCTAAGTATTAGGCGCTTGGATTACCGCCACTGCTCCAGAGGCTATGATG 1620
Qy 3754 AACTGTTGGAGCTGCCAGCCCTTACTCTTTTAACTTTTACTGCTTTTATATATAA 3813

Db 1621 AACTGTTGGAGCTGCCAGCCCTTACTGTTTAACTATTTCAGTCTTATATAA 1680
Qy 3814 AGAAACAACCTAACGA 3829
Db 1681 AGAAACAACCTAACGA 1696
RESULT 5
ACC83474
ID ACC83474 standard; cDNA; 1860 BP.
XX
XX ACC83474;
DT 08-SEP-2003 (first entry)
XX
DE Rat chondrocyte-derived zinc finger protein Czf-1 cDNA.
XX
XX Rat; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritic;
KW antirheumatic; gene; ss.
XX
XX Rattus sp.
XX
XX Key Location/Qualifiers
FT CDS 176..1801
FT /*tag= a
FT /product= "Rat Czf-1"
XX
XX W02003044159-A2.
XX
XX 30-MAY-2003.
XX
XX 20-NOV-2002; 2002WO-IL000925.
XX
XX 20-NOV-2001; 2001US-0331626P.
XX
XX (PROC-) PROCHON BIOTECH LTD.
XX
XX Yaron A, Blumenstein S, Harari D;
XX
XX WPI; 2003-457599/43.
DR P-PSDB; ABR42913.
XX
XX New chondrocyte-derived zinc finger polypeptides and encoding
PT polynucleotides, useful for detecting, diagnosing and treating Czf-1
PT protein-related diseases, such as osteoarthritis and rheumatoid
PT arthritis.
XX
XX Example 1; Fig 1; 64pp; English.
XX
XX The present sequence is that of cDNA encoding a novel rat zinc finger
CC protein, designated Czf-1. The cDNA was isolated from a glioma cDNA
CC library using a probe from the 3' end of the Czf-1 gene. Czf-1 is
CC expressed in a variety of tissue types, particularly osteoblasts and
CC chondrocytes. It is expressed in the nuclei of hypertrophic chondrocytes
CC of the growth plate, and exhibits an alternate expression and
CC intracellular localisation pattern in cells expressing high, i.e.
CC constitutively active, levels of fibroblast growth factor receptor
CC (FGFR). It is up-regulated in osteoarthritic cartilage tissue and serves
CC as a marker for osteoarthritis. Czf-1 has been found in rats, mice and
CC humans. Sequence identity for nucleotide and amino acid sequences is
CC above 85%. Czf-1 polynucleotides, polypeptides and antibodies can be used
CC in the characterisation, diagnosis and treatment of FGFR-related and
CC skeletal diseases and disorders, such as osteoarthritis, rheumatoid
CC arthritis, and cartilage related diseases
XX
XX Sequence 1860 BP; 366 A; 564 C; 601 G; 329 T; 0 U; 0 Other;
Query Match 39.1%; Score 1504.4; DB 7; Length 1860;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 72; Indels 3; Gaps 1;
Qy 1917 AGTGATACCTAGAGATTCCTTCAGCAAGGCTCCTGGGCCCTGGAGCGATT 1976

229 AGTGGATCTGAAGATTCTCTCGACGAAGGACCTGGGCTCTGGTGTGGAGATGATT 288
 1977 GCTACTAGCCCAAGATCTGGAGTTTAAAGAGAAAGAGAGATGAAGTGACGGCCA 2036
 289 GCTACTAGCCCAAGATCTGGAGTTTAAAGAGAAAGAGAGATGAAGTGACGGCCA 348
 2037 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC 2096
 349 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGAGCCAGACC 408
 2097 TGGAGCTTCCCTATGGGCTGAGTGACGACGAGTCTGGGGCGGCCGCGACATAAGTGCAGA 2156
 409 TGGAGCTTCCCTATGGGCTGAGTGACGATGAGTCTGGGGCGGCCGCGACATAAGTGCAGA 468
 2157 GAGTGAAGTTGAGAACCAAGCCAGGGGTCAGGGAGGCCAGGGGTGAGAGGCCAGGGCC 2216
 469 AAGTGAAGTTGAGAACCAAGCCAGGGGTCAGGGAGGCCAGGGGTGAGAGGCCAGGGCC 528
 2217 AGCCTGTGAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTGTGTGGGGCGAGAGGGCG 2276
 529 AGCCTGTGAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTGTGTGGGGCGAGAGGGCC 588
 2277 GGGTGGGGGGCCCCCTGCTGCCGCCACAGGTTACTGTAATCTGCTGCTGCTGCTGCTGCT 2336
 589 GGGTGGGGGGCCCCCTGCTGCCGCCACAGGTTACTGTAATCTGCTGCTGCTGCTGCTGCT 648
 2337 GTCCCACTACTGAGCCACTGAAGCGGCACATGACAGACACAGCGGGGAGAACCGTT 2396
 649 GTCCCACTACTGAGCCACTGAAGCGGCACATGACAGACACAGCGGGGAGAACCGTT 708
 2397 CGCTGTGCGGCTGCCCATACGCTGAGCCAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 2456
 709 CGCTGTGCGGCTGCCCATACGCTGAGCCAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 768
 2457 CACCCATCTGCGGAGAGCCCTACCGTTGCTCCCACTGCTCCCTGCTGCTGCTGCTGCTGCT 2516
 769 CACCCATCTGCGGAGAGCCCTACCGTTGCTCCCACTGCTCCCTGCTGCTGCTGCTGCTGCT 828
 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGCTCCCACTCTCTCTCTGCCCCAAC 2576
 829 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGCTCCCACTCTCTCTCTGCCCCAAC 888
 2577 CTGTGCTTTCGATCTGTGCTCCAGACCAACCGGCTCCCACTCCCACTCCCTGCCCCAAC 2636
 889 CTGTGCTTTCGATCTGTGCTCCAGACCAACCGGCTCCCACTCCCACTCCCTGCCCCAAC 948
 2637 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGCTCCCACTCTCTCTCTGCCCCAAC 2696
 949 AGGACCAATGCTCCCGGCTCAGAGATGCTCTCATCTCCCACTGAGTCTCCATCT 1008
 2697 GCACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2756
 1009 GCACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
 2757 CTGTGCTTTCGATCTGTGCTCCAGACCAACCGGCTCCCACTCCCACTCCCTGCCCCAAC 2816
 1069 CTGTGCTTTCGATCTGTGCTCCAGACCAACCGGCTCCCACTCCCACTCCCTGCCCCAAC 1128
 2817 CTGTGCAAGCAACTGAGAGAGGTGAGGACAGAGCTGGGAGCTGCCATGTGTGGGG 2876
 1129 CTGTGCAAGCAACTGAGAGAGGTGAGGACAGAGCTGGGAGCTGCCATGTGTGGGG 1188
 2877 CTGCATCGAGGAGAGCTGGAGGGTGGCATCTGGGGAGCCCGAGGCGCTGTCACAA 2936
 1189 CTGCATCGAGGAGAGCTGGAGGGTGGCATCTGGGGAGCCCGAGGCGCTGTCACAA 1248
 2937 AGGCTTTGCTGATGTTATGCTCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2996
 1249 AGGCTTTGCTGATGTTATGCTCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1308
 2997 CATGAAGCTCAGAGTGGTGAAGAACCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3056
 1309 CATGAAGCTCAGAGTGGTGAAGAACCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1368

QY 3057 TCATCTGGATAACCTGAACCGCACAGCGCGTCCACAGGAGAAAAGCCCTACAGTG 3116
 Db 1369 GCATCTGGATTAACCTGAACCGCACAGCGCGTCCACAGGAGAAAAGCCCTACAGTG 1428
 QY 3117 CCCCCTCTGTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
 Db 1429 CCCCCTCTGTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1488
 QY 3177 CTCTGGTGAACAAACCTTTTCGGTGTAGCCTTTTCGAACTACAGCTGCAACAGAGATGAA 3236
 Db 1489 CTCTGGTGAACAAACCTTTTCGGTGTAGCCTTTTCGAACTACAGCTGCAACAGAGATGAA 1548
 QY 3237 CCTCAACAGCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3296
 Db 1549 CCTCAACAGCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1608
 QY 3297 CTATACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3356
 Db 1609 TTACACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668
 QY 3357 AGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3416
 Db 1669 AGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728
 QY 3417 TGTCTTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3476
 Db 1729 TGTCTTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788
 QY 3477 AGACTCACTTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3536
 Db 1789 AGACTCACTTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
 QY 3537 TCCTGCACTTTTATAC 3551
 Db 1846 TCCTGCACTTTTATAC 1860

RESULT 6
 ABA82987
 ID ABA82987 standard; DNA; 1975 BP.
 XX
 AC ABA82987;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Human transcription factor TRFX-14 coding sequence.
 XX
 KW Human; transcription factor; TRFX; cell proliferative disease;
 KW autoimmune disease; inflammation; neurological disease;
 KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
 KW neuroprotective; antiinflammatory; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200172777-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 13-MAR-2001; 2001WO-US008117.
 XX
 PR 13-MAR-2000; 2000US-0188986P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hillman JL, Baughn MR, Yue H, Lal P, Lu DM, Patterson C;
 PI Azimzai Y, Bandman O, Tang YT, Machur P, Shah P, Au-Young J;
 PI Reddy R;
 XX
 DR WPI; 2001-570896/64.
 DR P-PSDB; ABB50163.
 XX
 PT Novel transcription factor polypeptides, used to treat diseases

PT associated with altered activity and expression of TRFX, and to screen
 PT for agents capable of modulating its activity.
 XX
 PS Claim 11; Page 251; 327pp; English.
 XX
 CC The present sequence is the coding sequence for a human transcription
 CC factor. The transcription factor and its coding sequence are useful in
 CC the diagnosis, treatment and prevention of diseases associated with
 CC altered expression of the transcription factor e.g. cell proliferative,
 CC autoimmune/inflammatory, neurological and developmental disorders. A
 CC number of specific disorders/diseases are given in the specification,
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
 CC allergies, anemia, asthma, autoimmune thyroiditis, bronchitis, atopic
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
 CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections
 XX
 SQ Sequence 1975 BP; 396 A; 625 C; 612 G; 342 T; 0 U; 0 Other;
 Query Match 38.9%; Score 1498; DB 5; Length 1975;
 Best Local Similarity 88.6%; Pred. No. 0;
 Matches 1707; Conservative 0; Mismatches 176; Indels 44; Gaps 6;
 QY 1917 AGTGATCTGAGTCTCTCGAGAGGTCCTGGGCGCCCTGCTGTGGAGCGGATTT 1976
 DB 86 AGTGATCTGAGTCTCTCGAGAGGTCCTGGGCGCCCTGCTGTGGAGCGGATTT 145
 QY 1977 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGGATGAAGTGACGCCA 2036
 DB 146 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGGAGAGAGCGACGCCA 205
 QY 2037 CAACACAGATCTGGGCTTTGAGAGAGATCTGAGAGAGATCTCAGGGGCGCAGACC 2096
 DB 206 CAGTACAGATCTGAGGCTTTGAGAGAGATCTGAGAGAGATCTCAGGGGCGCAGGCC 265
 QY 2097 TGGACTTCCCTATGGGCTGAGTGACAGAGTCTGGGCGCGCCGCACTAAGTGCGGA 2156
 DB 266 TGGGCTTCCCTATGGCTGAGCGAGATGAGTCTGGGCGCGCCGCACTAAGTGCGGA 325
 QY 2157 GAGTGAAGTTGAGAAACAGCAGGGGTCAGGGAGGCGCAGGGTGAGAGCGCAGGCC 2216
 DB 326 GAGTGAAGTTGAGAGGAGCAGCAGGGGTCAGGGAGGCGCAGGGTGAGAGCGCAGGCC 385
 QY 2217 AGCCTGTGAGTGTGGGGGCGCAGAGTGTGGGGGCGCGTGTGTGGGGCAGAGGCGG 2276
 DB 386 AGCCTGCCAGTGTGTGGGGGCGCAGAGTGTGGGGGCGCGTGTGTGGGGCAGAGGCGG 445
 QY 2277 GGGTGGGGGCGCCCGTGCCTCCCGCAGTGTGATGATGATGATGATGATGATGATGATG 2336
 DB 446 GGGTGGGGGCGCCCGTGCCTCCCGCAGTGTGATGATGATGATGATGATGATGATGATG 505
 QY 2337 GTCCCACTACTGAGCCACCTGAAGCGGCACATGACAGACACACAGCGGGAGAGCGGTT 2396
 DB 506 GTCCCACTACTGAGCCACCTGAAGCGGCACATGACAGACACACAGCGGGAGAGCGGTT 565
 QY 2397 CGCGTGTGGCGCTGCCCATACGNTACCGCCAGTTCCTCACTGAGCGGACATACCGG 2456
 DB 566 CGCGTGTGGCGCTGCCCATACGNTACCGCCAGTTCCTCACTGAGCGGACATACCGG 625
 QY 2457 CACCCATCTGAGGAGAGCGCTACCGTTGCTCCCACTGCGCCCTTGGCTCGCAGCGCT 2516
 DB 626 CACCCATCTGAGGAGAGCGCTACCGTTGCTCCCACTGCGCCCTTGGCTCGCAGCGCT 685
 QY 2517 GGGCAACCTGAGCGCGCATCAGCGCACCCACACAGGGGCTTCCCACTTCTCCCTGCCAAC 2576
 DB 686 GGGCAACCTGAGCGCGCATCAGCGCACCCACACAGGGGCTTCCCACTTCTCCCTGCCAAC 745
 QY 2577 CTGTGGCTTTCATGCTGTGCTCAGGACCAACCGGGCTTCCGATCCACAGAGCAGGA 2636
 DB 746 CTGTGGCTTTCATGCTGTGCTCAGGACCAACCGGGCTTCCGATCCACAGAGCAGGA 805

QY 2637 GGGGCAATGCCCCCAGCATCAGAAAAATGCGCTGATCCTGCCAGACTTGAGTCTTCAATGT 2696
 DB 806 GGGGCGGTGCCCCCGCGACCTGAAGATGCTCTGCTCTTCCAGATTTGAGCCTCATGT 865
 QY 2697 GCACCCAGTGTGCGAGTTTCTTCCAGAGCTGTGGGCGAGCTGCGGGTGAAGGGAGAG 2756
 DB 866 GCGCCAGTGTGCGAGTTTCTTCCAGAGCTGTGGGCGAGCTGCGGGTGAAGGGAGAG 925
 QY 2757 CTTGTGTGAATCGGATCCGAAACCACTGCGAGAGTACTGTTTCCCTTGGACCTGCGGG 2816
 DB 926 CTTGTGTGAATCGGATCCGAAACCACTGCGAGAGTACTGTTTCCCTTGGACCTGCGGG 985
 QY 2817 CTGTGACAGAGTGTGAGGAGGCTGAGGCGAGAGCTGCGGAGCTGCGCATGTGTGGCG 2876
 DB 986 CTGTGACAGAGTGTGAGGAGGCTGAGGCGAGAGTGTGCGGAGCTGCGCATGTGTGGCG 1045
 QY 2877 CTGCATGCGCAGAGAGGCTGAGGAGGCTTGCCTACTGCGGAGACCCAGGGGCTTGTGCA 2936
 DB 1046 CTGCATGCGCAGAGAGGCTGAGGAGGCTTGCCTACTGCGGAGACCCAGGGGCTTGTGCA 1105
 QY 2937 AGCTTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2996
 DB 1106 AGCTTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1165
 QY 2997 CATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTGCGCTGTGCGCTGTG 3056
 DB 1166 CATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTGCGCTGTGCGCTGTG 1225
 QY 3057 TCATCTGATTAACCTGAAACCGGACACAGCGCTTCCACACAGAGAGAAAGCCCTTCAAG 3116
 DB 1226 TCATCTGATTAACCTGAAACCGGACACAGCGCTTCCACACAGAGAGAAAGCCCTTCAAG 1285
 QY 3117 CCGCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3176
 DB 1286 CCGCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345
 QY 3177 CTCTGTGACAAACCTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATGAA 3236
 DB 1346 CTCTGTGACAAACCTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATGAA 1405
 QY 3237 CTTCAACCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3296
 DB 1406 CTTCAACCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1465
 QY 3297 CTATACCACAGGCGCTGCGGACCACTACAAGCTCATACAAGGTCATGAGGTCATGTTG 3356
 DB 1466 CTATACCACAGGCGCTGCGGACCACTACAAGCTCATACAAGGTCATGAGGTCATGTTG 1525
 QY 3357 AGCAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3416
 DB 1526 AGCAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1585
 QY 3417 TGTGTTGAGCCTGCGGGTCCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3476
 DB 1586 TGTGTTGAGCCTGCGGGTCCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1645
 QY 3477 AGACTCAGCTTGAACCTAAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3536
 DB 1646 AGACTCAGCTTGAACCTAAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672
 QY 3537 TCCTGCAATTTTACAAATGAACCTAGAAACCACTTCTCCCTTCTCCCTGCTGCTGCTG 3596
 DB 1673 TCCTGCAATTTTACAAATGAACCTAGAAACCACTTCTCCCTTCTCCCTGCTGCTGCTG 1728
 QY 3597 GGTCCACACAGACTAAGCTAGGCACTATATGAGCAGCTGAAATCCCATGTGTGAGGGG 3656
 DB 1729 GGTCCACACAGACTAAGCTAGGCACTATATGAGCAGCTGAAATCCCATGTGTGAGGGG 1788
 QY 3657 -CCATATAGACAGGGGA-CTTGTCTTAGCTACGTACAGATGAGCTAAGTATGAGG 3714
 DB 1789 CCATATGACAGGGGACCTTGTGCTTGTAGGACCTTACAGAGCTCAGTGAAGAGG 1848
 QY 3715 CTTTGAATTCACCGCACTGCTCCAGAGGCTATGAGTGAATGAACTGGTGTGGAG-CTGCCCA 3773

Db 1849 CCTGTATTCACCTCACTGCCCCAGGGGCTGTGACAAACCGGCTGGGGAGTGCCCA 1908
QY 3774 GCCTTTTAC-----TGTTTTAACCTTATTCAGTCTTTTATATAAAGGAAACACTAACAGA 3829
Db 1909 GCCTCCACCTGTTTATTTAACTTATTCAGTCTTTTATATAAAGGAAACACTAACAAA 1968
QY 3830 AAAAAAA 3836
Db 1969 AAAAAAA 1975
RESULT 7
ACC83473
ID ACC83473 standard; cDNA; 1855 BP.
XX
AC ACC83473;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human chondrocyte-derived zinc finger protein Czf-1 cDNA.
XX
KW Human; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritis;
XX
KW antirheumatic; chromosome 2p22-23; gene; 88.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
XX CDS 171..1796
XX FT /*tag= a
XX FT /product= "Human Czf-1"
XX
XX WO2003044159-A2.
XX
XX 30-MAY-2003.
XX
XX 20-NOV-2002; 2002WO-IL000925.
XX
XX 20-NOV-2001; 2001US-0331626P.
XX
XX (PROC-) PROCHON BIOTECH LTD.
XX
XX Yayon A, Blumenstein S, Harari D;
XX WPI; 2003-457599/43.
XX P-PSDB; ABR42912.
XX
XX New chondrocyte-derived zinc finger polypeptides and encoding
XX PT polynucleotides, useful for detecting, diagnosing and treating Czf-1
XX PT protein-related diseases, such as osteoarthritis and rheumatoid
XX PT arthritis.
XX
XX Claim 12; Fig 1; 64pp; English.
XX
XX The present sequence is that of cDNA encoding a novel human zinc finger
XX protein, designated Czf-1. Czf-1 is expressed in a variety of tissue
XX types, particularly osteoblasts and chondrocytes. It is expressed in the
XX nuclei of hypertrophic chondrocytes of the growth plate, and exhibits an
XX alternate expression and intracellular localisation pattern in cells
XX expressing high, i.e. constitutively active, levels of fibroblast growth
XX factor receptor (FGFR). It is up-regulated in osteoarthritic cartilage
XX tissue and serves as a marker for osteoarthritis. Czf-1 has been found in
XX rats, mice and humans. Sequence identity for nucleotide and amino acid
XX sequences is above 85%. The human gene maps to chromosome 2p22-23. Czf-1
XX polynucleotides, polypeptides and antibodies can be used in the
XX characterisation, diagnosis and treatment of FGFR-related and skeletal
XX diseases and disorders, such as osteoarthritis, rheumatoid arthritis, and
XX cartilage related diseases
XX
XX Sequence 1855 BP; 344 A; 590 C; 619 G; 302 T; 0 U; 0 Other;
Query March 36.2%; Score 1392; DB 7; Length 1855;
Best Local Similarity 92.6%; Pred. No. 0;

Matches 1461; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 1917 AGTGGATCTAGAGATTCTTTGACGAAGTCTCTGGGGCCCTGTTGTGAGAGCAGATT 1976
Db 224 AGTGGATCTAGAGATCTCTCGACGAGACCCGGGGCCCTGGTATTGAGAGATGATT 283
QY 1977 GTACTAGGCCAAGATCTGAGTTTGAAGAGAGAGAGAGAGATGAGGTGACGCCCA 2036
Db 284 GCTACTAGGCCAGGATCTGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGCA 343
QY 2037 CAACGACAGCTCATGGGCTTTGAGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACC 2096
Db 344 CAGTGACAGCTCATGGGCTTTGAGAGAGACTCTGGAAGAGAGACTCTCTGGGGCCAGGCC 403
QY 2097 TGGACTTCCCTATGGGCTGAGTGACGACGAGTCTGGGGGGCGCGCGCACTAAGTGGGGA 2156
Db 404 TGGGCTTCCCTATGGGCTGAGCGACATGAGTCTGGGGGGCGCGCGCACTAAGTGGGGA 463
QY 2157 GAGTGAAGTTGAGGAACCCAGGGGTCCAGGGGGAGCCAGGGGTGAGAGGGCCAGGCC 2216
Db 464 GAGTGAAGTTGAGGAGCCAGCAGGGGTCCAGGGGGAGCCAGGGGTGAGAGGGCCAGGCC 523
QY 2217 AGCCTGTCTGAGTGTGGGGGGCGGACAGGTGAGGGGCCGTGTTGTGGGGCAGAGGGGG 2276
Db 524 AGCCTGTCTGAGTGTGGGGGGCGGACAGGTGAGGGGGCGGTGTTGTGGGGCAGAGGGGG 583
QY 2277 GGGTGG 2336
Db 584 GGGTGG 643
QY 2337 GTCCCACTACTCTGAGCCACCTGGAAGCGGCACATGACAGACACAGCGGGGAGAGCCGTT 2396
Db 644 GTCCCACTACTCTGAGCCACCTGGAAGCGGCACATGACAGACACAGCGGGGAGAGCCGTT 703
QY 2397 CCGTGTGG 2456
Db 704 CCGTGTGG 763
QY 2457 CACCCATCTAGTGGCGAGAGCCCTACCGTTGTCCCACTGTCCCGCTTTCCTGCGACAGCCT 2516
Db 764 CACCCATCTAGTGGCGAGAGCCCTACCGTTGTCCCACTGTCCCGCTTTCCTGCGACAGCCT 823
QY 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGGCTCCCACTCTCTCTCCGCCCAAC 2576
Db 824 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGGCTCCCACTCTCTCTCCGCCCAAC 883
QY 2577 CTGTGGCTTTGATGCTGTCTCCAGACCAACCCGGCTCCCACTGTCACAGAGCAGGA 2636
Db 884 CTGTGGCTTTGATGCTGTCTCCAGACCAACCCGGCTCCCACTGTCACAGAGCAGGA 943
QY 2637 GGGGACAAATGCCCGACGATCAGAAATGCGCTGATCTCTGCCAGACTTGAGTCTTCATGT 2696
Db 944 GGGGGCGGTGCGCGGGGACCTGAAGATGCTCTCTCTCTCCAGATTTGAGCCTCCATGT 1003
QY 2697 GCCACAGGTGGTCCAGTTTCTGCGCAGACTGTGGGAGCTGCGGGGTGAGGGGAGAG 2756
Db 1004 GCCACAGGTGGTCCAGTTTCTGCGCAGACTGTGGGAGCTGCGGGGTGAGGGGAGAG 1063
QY 2757 CTGTGTGGAATCGATCCGAACCACTGCCAGAGCTACTGTTCCTCTTGAGACTCCCGGG 2816
Db 1064 CTCTGCGGACTGTGATCAGAACCACTGCCAGAGCTGCTATTCTCTTGAGACTCCCGGG 1123
QY 2817 CTGTGACAGAACTGAGAGAGGTGAGGGCAGCAGAGCTGGAGCTGCGCATGTGTGGCG 2876
Db 1124 CTGTGACAGAGCTGAGAGAGGTGAGGGGTAGTTCGGCTGGGAGCTGCCATGTGTGGGCG 1183
QY 2877 CTGCATGCGAGAGAGGCTGAGGGGGTTCACCTGCGGGGACCCCGAGGGCCCTGGTGACAA 2936
Db 1184 CTGCATGCGAGAGAGGCTGAGGGGGTTCACCTGCGGGGACCCCGAGGGCCCTGGTGACAA 1243
QY 2937 AGGCTTTGCTGTAGTTTATGCCCTTTTGCACCTCATCTACCCCAACACCTGGCTCGGCA 2996
Db 1244 AGGCTTTGCTGTAGCTCTGCCCTTTTGCACCTCACTATCCCAACACCTGGCTCGGCA 1303

PD		24-AUG-2000.	
XX		18-FEB-2000; 200WO-US004340.	
PP		19-FEB-1999; 99US-0120680P.	
XX		23-APR-1999; 99US-00298733.	
PR		17-AUG-1999; 99US-0149639P.	
PR		23-SEP-1999; 99US-0155696P.	
PR		01-OCT-1999; 99US-0157247P.	
PR		29-NOV-1999; 99US-0167822P.	
PR		29-NOV-1999; 99US-0167823P.	
PR		15-FEB-2000; 2000US-0182711P.	
XX		(ALPH-) ALPHAGENE INC.	
PA		Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;	
PPI		WPI; 2000-549267/50.	
XX		P-PSDB; AAB23631.	
DR		New secreted proteins and polynucleotides encoding them, which are	
PT		derived from Homosapiens, useful for therapy, diagnosis, and research, as	
PT		well as nutritional sources or supplements.	
XX		Claim 70; Page 288-289; 309pp; English.	
PS		The present sequence is the coding sequence for a human secreted protein.	
XX		The sequence was isolated from an adult lung cDNA library. The proteins	
CC		and coding sequences of the invention can be used in the isolation of	
CC		similar genes and proteins, in the elucidation of their function in vivo,	
CC		and to treat a number of conditions. It is possible that they may have	
CC		uses as nutritional supplements, as cytokine or cell proliferation	
CC		factors, in immune modulation, where they may be used to treat immune and	
CC		autoimmune diseases, as haematopoiesis regulators (treating myeloid or	
CC		lymphoid cell deficiencies), in the promotion of tissue growth, they may	
CC		have chemokine or chemotactic activity, haemostatic or thrombolytic	
CC		activity, or anti-inflammatory activity	
XX			
SQ		Sequence 1996 BP; 449 A; 605 C; 578 G; 364 T; 0 U; 0 Other;	
	Query Match	35.5%; Score 1367.6; DB 3; Length 1996;	
	Best Local Similarity	86.2%; Pred. No. 0;	
	Matches 1598; Conservative	0; Mismatches 210; Indels 45; Gaps 6;	
QY	2003	AAGAGGAGAGGACAGCAGATGAAGTGACGCCACACAGCACAGCTCATGGCTTTGAGA	2062
Db	139	AAAGTAAAAGGGGAGAGTTGGAGCGCTGATCCTACCCTATGCTGATGTCTTCTTTATGT	198
QY	2063	GAGACTCTGAAGAGACACTCTCAGGGGGGCAGACCTGGACTTCCTATGGGCTGATGTACG	2122
Db	199	CTATTTCACGAGACTCTCTGGGGGCCAGGCTTCCTATGGGCTGAGCGACG	258
QY	2123	ACGAGTCTGGGGCGGCGCGCACTAAGTGCAGAGGTGAAGTTAGGAACACGACGAGG	2182
Db	259	ATGAGTCTGGGGCGGCGCGGACCTAAGTGCAGAGGTGAAGTTAGGAGACGACGACGAGG	318
QY	2183	GTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGAGTGTGGGGGGCGCA	2242
Db	319	GTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGAGTGTGGGGGGCGCA	378
QY	2243	CAGGTGAAGGGCGCGTGTGTGGGCGAGAGGGCGGGGCCCCCGCTGCCCCCAC	2302
Db	379	CAGGTGAAGGGCGCGTGTGTGGGCGAGAGGGCGGGGCCCCCGCTGCCCCCAC	438
QY	2303	GGTTACTGTACTCATGCCGCTGTGCGCTTTCTGTCTCCACTACTCTCGAGCCACCTGAAGC	2362
Db	439	GGCTACTGTACTCATGCCGCTCTGCACCTTCTGTGTCCACTACTCTCGAGCCACCTGAAGC	498
QY	2363	GGCACAATGCAGACACACAGCGGGGAGAACCGCTTCGCTGTGGCGCGCTGCCCATACGNT	2422
Db	499	GGCACAATGCAGACACACAGCGGGGAGAACCGCTTCGCTGTGGCGCGCTGCCCATACGCT	558
QY	2423	CGAGCCAGTTCTGTCAACCTGACGCGGACATACCCCGCACCTACTCTGCGAGAGCCCTACC	2482

Db 1628 -----TAGGTCTCTTCTCCCATGTTTTATACAGACGGACCAG 1666

Qy 3563 AAACCACTTTCCCTTTCTCCCCCGTGTGTACGGGCTCCACACAGACTAACCTTAGGCAC 3622

Db 1666 AAGCCACCTT-----TTTCTCCCGCTGTGGCCAGGGGCTCCACACAGACTAACCTTAGGCAC 1721

Qy 3623 TATATGACCAAGCCTGNAATCCCATGCTCAGGGGG-CCATATAGACCAAGGGGA-CTTGTCT 3680

Db 1722 TATAAGACCAAGCCCAACCCCATCTGGGGGGGGGGCCCATATGACCAAGGGGACCTTGCCT 1781

Qy 3681 TAGCTCAGCTACCAAGATGAGCTAAGTGATTAGGGCCTTGGATTACCCGCCACTCTCCCA 3740

Db 1782 TGACTGGGCACCTTCAGGAGCTCAGTGAGAAAGGCCCTGTATTCACTCCACTGCCCCCA 1841

Qy 3741 GAGCTATGATGAACACTGTGTGGAG-CTGCCACAGCCTTTTAC-----TGTTTAACTTA 3799

Db 1842 GGGGCTGTGGACAAACCGGCTGGGGGACTGCCAGCCTCCACCTGTTATTATTAACTTAT 1901

Qy 3795 TTTCAAGTCTTTTATAATAAAGGAAACACTAACAGAAAAAATAAAAAAAAAAAAA 3847

Db 1902 TTCAAGTCTTTTATAATAAAGGAAACACTAACAGAAAAAATAAAAAAAAAAAAA 1954

RESULT 9

AAI59942

ID AAI59942 standard; cDNA; 1488 BP.

XX AAI59942;

AC AAI59942;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 3931.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia; ss.

XX Homo sapiens.

OS WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-0048725.

PR 25-APR-2000; 2000US-0052317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao OA;

PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR P-PSDB; AAM40786.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Claim 1; SEQ ID NO 3931; 10078pp; English.

PS The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 3931; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC

encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Sequence 1488 BP: 287 A; 460 C; 484 G; 257 T; 0 U; 0 Other;

Query Match 31.7%; Score 1220.2; DB 4; Length 1488;

Best Local Similarity 92.9%; Pred. No. 4.1e-301;

Matches 1272; Conservative	0; Mismatches	97; Indels	0; Gaps	0;
----------------------------	---------------	------------	---------	----

QY	1885	GAGGGAGATTCCA	CAAGCACCTCCCTCTTTAGTGGATACTGAAGATTTCCTTCGACGAA	1944
DB	119	GAGGGAGAATCCA	CAAGCCCCCTCCCTCTTCAGTGGATACTGAAGACTCCCTCGACGAA	178
QY	1945	GGTCCTGGGGCCCTGGTGTGTGGAGAGCGATTTCTACTAGGCCAAGATCTGGAGTTTGAA	2004	
DB	179	GGACCCGGGGCCCTGGTATTGGAGAGTGATTTCTACTAGGCCAGGATCTGGAGTTTGAG	238	
QY	2005	GAGGAAGAGAA	GAGGANTGAAGTGACGGCCACAACGACCATCTATGGGCTTTGAGAGA	2064
DB	239	GAGGAAGAGAA	GAGGAGGAGGCGCGCAACAGTGACCCAGCTCATGGGCTTCGAGAGA	298
QY	2065	GACTCTGAAGGACATCTCAGGGGGCCAGACCTCTGGACTTCCCTATGGGCTGAGTGACGAC	2124	
DB	299	GACTCGAAGGAGACTCTCTGGGGGCCAGGCCCTTGGGCTTCCCTATGGGCTGACGACGAT	358	
QY	2125	GACTCTGGGGGGCGGCCCGCGCACTAAGTGCAGAGAGTGAAGTTGAGAA	CAAGCCACGGGGT	2184
DB	359	GAGTCTGGGGGGCGGCCGGGCACTAAGTGCAGAGAGTGAAGTTGAGGAGCCAGCACGGGGT	418	
QY	2185	CCAGGGAGGCCCAGGGGTGAGAGCCAGGCCCAGCCTGTCTGATGTGTGGGGGGCCGACA	2244	
DB	419	CCAGGGAGGCCCAGGGGTGAGAGGCCAGGCCAGCCTGCCAGCTGTGTGGGGGGCCGACA	478	
QY	2245	GGTGAAGGGGCCGTGTGTGTGGGGGAGAGGGCGCGGGGTGGGGGGCCCCCGCTGCCCCCCCACGG	2304	
DB	479	GGTGAAGGGGCCGTGTGTGTGGGGGAGAGGGCGCGGGGTGGGGGGCCCCCTGTCTGCCCCCCCACGG	538	
QY	2305	TTACTGTACTCATGSCCGGCTGTGGCTTTTCGTGTCCCACTACTCTGAGCCACCTCTGAAGCGG	2364	
DB	539	CTACTGTACTCATGCCGCCCTCTGCACCTTGTGTCCCACTACTCTGAGGCCACCTCTGAAGCGG	598	
QY	2365	CACATGAGACACACAGCGGGGAGAGCCCTTCCGTGTGGCCGCTGCCCATACGNNCTCA	2424	
DB	599	CACATGAGACACACAGCGGAGAGAGCCCTTCCGTGTGGCCGCTGCCCATACGCTCA	658	
QY	2425	GCCAGATTGCTCAACCTGACGACATACCGCACCCATACTGGCGAGAGAGCCCTACCGT	2484	
DB	659	GCCAGTCTGTCAACCTGACACGACATACCGCACCCATACTGGCGAGAGAGCCCTACCGC	718	
QY	2485	TGTCCTCCACTGCCCTTTTGCTGTGAGCAGCCTGGGCAACTGAGGGGGCATGACGGCAAC	2544	
DB	719	TGTCCTCCACTGCCCTTTTGCTGTGAGCAGCCTGGGCAACTGAGGGGGCATACGCGTACC	778	
QY	2545	CACACAGGGGCTCCCACTCTCTCCCTGCCCAACCTGTGGCTTTGATGCTCTGTCTCCACGA	2604	
DB	779	CACGAGGGGCCCCCACTCTCTCTGCCCAACCTGTGGCTTTGCTGCTGTATCTCCACGA	838	
QY	2605	CCAAACCGGGCTCCCACTCCACAGACGAGGAGGGGACAAATGCCCCGACGATCAGAAAAT	2664	
DB	839	CCAGCCGGGCTCCCACTCTCCACAGACGAGGAGGGGGCGGTGCCCCGGCACCTGAGAT	898	

PR 07-JUL-2000; 2000US-02116647P.
PR 07-JUL-2000; 2000US-02116800P.
PR 11-JUL-2000; 2000US-02117480P.
PR 11-JUL-2000; 2000US-02117487P.
PR 11-JUL-2000; 2000US-02117496P.
PR 14-JUL-2000; 2000US-02182900P.
PR 26-JUL-2000; 2000US-02202663P.
PR 26-JUL-2000; 2000US-02209664P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-02252113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256717P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 40206; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK54702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)

CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (i) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 9266 BP; 2324 A; 2431 C; 2502 G; 2009 T; 0 U; 0 Other;

Query Match 31.1%; Score 1195; DB 4; Length 9266;
 Best Local Similarity 81.3%; Pred. No. 2.9e-294;
 Matches 1568; Conservative 0; Mismatches 216; Indels 144; Gaps 10;

QY 2003 AAGAGGAGAGAGAGAGATGAGGTGACGCGCCACACAGCCAGCTCATGGCTTTGAGA 2062
 DB 1441 AAGTGAAGAGGAGAGATTGAGCCTGATCCTACCTATGCTGATGCTCTTTATGT 1500
 QY 2063 GAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACTGGACTTCCCTATGGGTGAGTGACG 2122
 DB 1501 CTATTTACAGGAGACTCTCTGGGGGCCAGGCTTGGCTTCCCTATGGGTGAGCGACG 1560
 QY 2123 ACGAGTCTGGGGGGCCGCCACTAAGTGGGAGAGTGAAGTTGAGGAACCGCCAGGG 2182
 DB 1561 ATGAGTCTGGGGGGCCGCCACTAAGTGGGAGAGTGAAGTTGAGGAGCCAGCCAGGG 1620
 QY 2183 GTCCAGGGGAGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGAGTGTGGGGGGCCGA 2242
 DB 1621 GTCCAGGGGAGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGAGTGTGGGGGGCCGA 1680
 QY 2243 CAGGTGAGGGGGCCGTGTTGGGGCAGAGGGCGGGTGGGGGGCCCGCTGCCGCCAC 2302
 DB 1681 CAGGTGAGGGGGCCGTGTTGGGGCAGAGGGCGGGTGGGGGG-CCCTGTGTCGCCCCAC 1739
 QY 2303 GGTACTGTACTCATGCGCGGTGTCGGTTCGTTGCCACTACTTCGAGCCACCTGGAAGC 2362
 DB 1740 GGTACTGTACTCATGCGCGCTTGCACCTTCGTTCCTACTCTCGAGCCACCTGAAGC 1799
 QY 2363 GGCATGACAGACACAGAGGGGAGAGCCGTTCGGTGTGGCGGTGCCCTACCGNT 2422
 DB 1800 GGCATGACAGACACAGAGGGGAGAGCCGTTCGGTGTGGCGGTGCCCTACCGCT 1859
 QY 2423 CAGCCAGTTCGTCAACCTGACCGCACATACCGCACCCACTACTGGCGAGAGCCCTACC 2482
 DB 1860 CAGCCAGTTCGTCAACCTGACACGACATACCGCACCCCACTGGCGAGAGCCCTACC 1919
 QY 2483 GTTGTCCCACTGCCCTTTGCTGCAGAGCCTGGGCAACTGAGGGCGCATCAGCGCA 2542
 DB 1920 GCTGTCCCACTGCCCTTTGCTGCAGAGCCTGGGCAACTGAGGGCGCATCAGCGTA 1979
 QY 2543 CCCACAGGGCTCCCACTCCCTCCCTGCCAACCTGGCTTTCGATGCTGTCTCCAC 2602
 DB 1980 CCCACAGGGCTCCCACTCCCTCCCTGCCAACCTGGCTTTCGCTGTCTCTCCAC 2039
 QY 2603 GACCAACCCGGCTCCCACTCCCAAGAGAGAGGGGAGCAATGCCCG- 2651
 DB 2040 GACCAACCCGGCTCCCACTCCCAAGAGAGAGGGGCGGGCCCGCGACCTGGAAG 2099
 QY 2652 - 2652
 DB 2100 TAAGACACACAGGACCAAGATCTTGGGACATGGGTGGGTGACCTTAGGAATGCTTGG 2159
 QY 2652 - 2652
 DB 2160 ATTGGATTATAGCCAGGCTTTGTGCCCAAGATGCTCTGCTCTCTCCAGATTGAGC 2219
 QY 2689 CTTTCATGTGCCACAGGTGGTGGCTTTCCTGCCAGACTGTGGCAGCTGGGGGTGAA 2748
 DB 2220 CTTTCATGTGCCACAGGT-GTCCAGATTTCCTGCCAGACTGTGGGCGAGCTGGGGGTGAA 2278
 QY 2749 GGGGAGAGCTGTGTGGAACCTGGATCCGAACCACTGCCAGAGCTACTGTTCCTTGGACC 2808

DB 2279 GGGGAGGGGCTCTGCGGAGCTGGATCAGAAACCACTGCCAGAGCTGCTATTCCCTTGGACC 2338
 QY 2809 TCCCGGGCTGTGGACAGGAACTGGAGGAGGTGAGGCGAGCAGGCTCGGAGCTGCCATG 2868
 DB 2339 TCCCGGGCTGTGGACAGGAACTGGAGGAGGTGAGGCTGCTGGCTGGAGCTGCCATG 2398
 QY 2869 TGTGGCGCTGTGATCCGAGGAGAGCTGGAGGGTGGCACTGTGGGGAGCCCGAGGGCCCT 2928
 DB 2399 TGTGGCGCTGTGATCCGAGGAGAGCTGGAGGGGTGCGAGTGGGGGGCCACACAGCTGT 2458
 QY 2929 GGTGACAAAGGCTTTGCTGTAGTTTATGCCCTTTGCACTACCTACCCCAACACCTG 2988
 DB 2459 - - - - - ACAAAGGCTTTGCTGTAGCTTGTGCCCTTTGCCACTTCTATCCCAACACCTG 2514
 QY 2989 GCTTGGCCATGAAGACTCACAGTGTGTGAGAAACCTTCCGCTGTGCCGCTGTCCATAC 3048
 DB 2515 GCCCGGCACATGAAGACACACAGTGTGTGAGAAAGCTTCCGCTGTGCCGCTGTCTTAT 2574
 QY 3049 GCTTGTGCTCATCTGGATTAACCTGAAACGGCACAGCGGTCCACACAGGAGAAAGGCC 3108
 DB 2575 GCTTGTGCTCATCTGGATTAACCTGAAACGGCACAGCGGTCCACACAGGAGAGAGGCC 2634
 QY 3109 TACAAGTGGCCCTCTGTCCGTATGCTGTGCAACCTTGGCCAACTCAAGCGTCTATGT 3168
 DB 2635 TACAAGTGGCCCTCTGTCCCTTATGCTGTGCAATCTGGCCAACTCAAGCGTCTATGT 2694
 QY 3169 CGCATCCACTCTGTGTGACAAACCTTTCGCTGTAGCTTGTGCAACTACAGCTGCAACAG 3228
 DB 2695 CGCATCCACTCTGTGTGACAAACCTTTCGCTGTAGCTTGTGCAACTACAGCTGCAACAG 2754
 QY 3229 AGTATCAACCTCAACGCTCATATGCTGCGACACAGCGGCGAGAACCTTCCGCTGTGTC 3288
 DB 2755 AGCATGAACCTCAACGCTCATATGCTGCGGACACAGCGGCGAGAACCTTCCGCTGTGTC 2814
 QY 3289 ACTGCGCTTATACACAGGCCACTTGGGCAACACTACAGCGTCTACAGAGGTGATGTC 3348
 DB 2815 ACTGCGCTTATACACAGGCCACTTGGGCAACACTACAGCGTCTACAGAGGTGATGTC 2874
 QY 3349 CATGTGGAGAGCAGGGGCTGTGCTCTCTGCCCTGAGGGCTGGGCCCACTCATATAGC 3408
 DB 2875 CACGTGGGGCAGAGGGCTGTGCTCTCTGCTGTGAGGGTGGGCCCACTCATATAGC 2934
 QY 3409 CCACTCTGTGTTTGTAGCACTGGGCTCAGCAGGCCCTGGGTGTCTAGGTAGCAGGGCT 3468
 DB 2935 CCACTCTGTGTTTGTAGCACTGGGCTCAGCAGGCCCTGGGTGTCTAGGTAGCAGGGCT 2994
 QY 3469 CTTTATTCAGACTCAGCTTGAACCTAAGTCTTCTTTTACCTGGGCTCTAGGAATTAGC 3528
 DB 2995 GTCCACAGACTCATCTGAAC- 3021
 QY 3529 CCTATGCTCTCTGCTATTTTATACAAATGAACCTAGAAACCACTTTCCTTTCTCCCGCC 3588
 DB 3022 TCCTTCTTCCCATGTTTATACAGAGCGACACAGAACCACTT- - - - - TTTCTCCCGCC 3077
 QY 3589 TGGTACAGGGCTCCACACAGACTAAGCTAGGCACTATATGGACCACTGAAATCCCATGG 3648
 DB 3078 TGGCCACAGGGCTCCACACAGACTAAGCTAGGCACTATATAGGACCACTGAAATCCCATGG 3137
 QY 3649 TCAGGGGG-CCATATAGACCAAGGGG-CTTGTCTTAGCTCAGCTACAGTACAGTACAGTAACT 3706
 DB 3138 CGGGGGGGCCCATATGACCAAGGGGACCTTGCCTTACTGAGGCACTTACAGAGCTCAGT 3197
 QY 3707 GATTAGGGCTTGGATTACCGCCACTGTCTCCAGAGGCTATGGATGAATGTTGGAG 3766
 DB 3198 GAGAAGGGGCTGTATTCACCTCCCTGCCCGGGCTGTGGACAAACCCGCTGGGG 3257
 QY 3767 -CTGCCCAGGCTTTTAC- - - - - TGTTTTAACTTATTTTCACTGCTTTTATAAAGGAAACA 3821
 DB 3258 ACTGCCAGGCTCCACCTGTTTATTACTTATTTTCACTGCTTTTATAAAGGAAACA 3317
 QY 3822 CTAAACAGA 3829
 DB 3318 CTAACAAA 3325

RESULT 11
AAL04937
ID AAL04937 standard; DNA; 9266 BP.
XX AC AAL04937;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 7625.
XX DE Human reproductive system related antigen; reproductive system disorder;
XX KW Human, reproductive system related antigen; cancer; gene therapy; ds.
XX KW cancer; gene therapy; ds.
XX OS Homo sapiens.
XX PN W0200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001339.
XX PF 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214866P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 14-AUG-2000; 2000US-0225759P.
XX PR 18-AUG-2000; 2000US-0226273P.
XX PR 22-AUG-2000; 2000US-0226681P.
XX PR 22-AUG-2000; 2000US-0226686P.
XX PR 22-AUG-2000; 2000US-0227182P.
XX PR 23-AUG-2000; 2000US-0227009P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 05-SEP-2000; 2000US-0229513P.
XX PR 06-SEP-2000; 2000US-0230437P.
XX PR 06-SEP-2000; 2000US-0230438P.
XX PR 08-SEP-2000; 2000US-0231243P.
XX PR 08-SEP-2000; 2000US-0231244P.
XX PR 08-SEP-2000; 2000US-0231413P.
XX PR 08-SEP-2000; 2000US-0231414P.
XX PR 08-SEP-2000; 2000US-0233080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 7625; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX Sequence 9266 BP; 2324 A; 2431 C; 2502 G; 2009 T; 0 U; 0 Other;
SQ
Query Match 31.1%; Score 1195; DB 4; Length 9266;
Best Local Similarity 81.3%; Pred. No. 2.9e-294;
Matches 1568; Conservative 0; Mismatches 216; Indels 144; Gaps 10;
QY 2003 AAGAGGAGGAGGAGGATGAAGTGAGCGGCACACAGCAGCTCATGGCTTTGAGA 2062
DB 1441 AAGTGAAAGGGAGGAGTTGGAGCTGATCTACCTATGCTGATCTCTTCTATGT 1500
QY 2063 GAGACTCTGAAGAGACTCTCAGGGGCGCAGACCTGGACTTCCCTATGGCTGAGTGACG 2122
DB 1501 CTATTTTACAGGAGACTCTCTGGGGCGCAGGCTTGGGCTTCCCTATGGCTGAGCGACG 1560
QY 2123 ACGAGTCTGGGGCGCGCCGCACTAAGTGCGGAGAGTCAAGTTGAGGAACAGCAGCGG 2182
DB 1561 ATGAGTCTGGGGCGCGCCGCGCACTAAGTGCGGAGAGTCAAGTTGAGGAGCCAGCGGG 1620
QY 2183 GTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCCAGCCCTGTGAGTGTGGGGGGCGGA 2242
DB 1621 GTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCCAGCCCTGTGAGTGTGGGGGGCGGA 1680
QY 2243 CAGGTGAGGGGCGGTGTTGTGGGCGAGGAGGGGGGTGGGGGCGCCGCTGCCCGCCAC 2302
DB 1681 CAGGTGAGGGGCGGTGTTGTGGGCGAGGAGGGCGGGTGGGGG-CCCTGTGCCCGCCAC 1739
QY 2303 GGTACTGTACTCATGCCGCTGTGGCTTTGCTGTCCCACTACTCGAGCCACCTGAAGC 2362
DB 1740 GGCTACTGTACTCATGCCGCTGTGACCTTGTGTCCCACTACTCGAGCCACCTGAAGC 1799
QY 2363 GGCACTGACAGACACAGCGGGGAGAGCCGTTCCGCTGTGGCGCTGCCCATACGNT 2422
DB 1800 GGCACTGACAGACACAGCGGGGAGAGCCGTTCCGCTGTGGCGCTGCCCATACGNT 1859
QY 2423 CAGCCAGTTCGTCAACCTGACCGCAGATACCGCACCTACTGTGGCGAGAGCCCTTACC 2482
DB 1860 CAGCCAGTTCGTCAACCTGACCGCAGATACCGCACCTACTGTGGCGAGAGCCCTTACC 1919
QY 2483 GTTGTCCCACTGCCCTTGTCCCTGACGAGCTGGGCAACCTGAGCGGAGCTCAGCGCA 2542
DB 1920 GCTGTCCCACTGCCCTTGTCCCTGACGAGCTGGGCAACCTGAGCGGAGCTCAGCGCA 1979

QY 2543 CCCACACAGGGGCTCCCACTCTCTCCCTGCCCAACCTGTGGCTTTGATGTGTGCTCCAC 2602
DB 1980 CCCACGACGGGCCCCCACTCTCTCCCTGCCCAACCTGTGGCTTTGATGTGTGCTCCAC 2039
QY 2603 GACCAACCGGGCTCCCACTCTCCCTGCCCAACCTGTGGCTTTGATGTGTGCTCCAC 2651
DB 2040 GACCAACCGGGCTCCCACTCTCCCTGCCCAACCTGTGGCTTTGATGTGTGCTCCAC 2099
QY 2652 ----- 2651
DB 2100 TAAGACACACAGGGGACCAAGATCTTGGGACATGGGTGGCTGACCTTAGGAATCTTGG 2159
QY 2652 -----ACGATCAGAAATGGCTGATCTCTGCTCTCTTCCAGATTTGAGT 2688
DB 2160 ATTGGATTATAGCCAGGTCTTTGTCCCAAGATGCTCTCTCTTCCAGATTTGAGC 2219
QY 2689 CTTTATGTCCCAACAGGTGTCCTGATTTCTGTCAGACTGTGGGACGTGCGGCTGAA 2748
DB 2220 CTCATGTGCCACAGGT-GTGCCAGTTTCTTCCAGACTGTGGGACGTGCGGCTGAA 2278
QY 2749 GGGGAGAGCTTGTGGAATCGGATCCGACACTGTCAGAGCTACTGTTCCTCTTGGACC 2808
DB 2279 GGGGAGGGGCTCTGGGGGACTGGATCAGAAACACTGTCAGAGCTACTTATTCCTTTGGACC 2338
QY 2809 TGCCGGGGCTGTGGACAGGAACTGGAGGAGGTGAGGGCAGCAGGCTGGGAGCTGCCATG 2868
DB 2339 TGCCGGGGCTGTGGACAGAGCTGGAGGAGGTGAGGGTAGTCCGCTGGGAGCTGCCATG 2398
QY 2869 TGTGGGCGCTGATGCGAGAGAGGTGAGGGGTGTCACCTGGGGGACCCAGAGGCGCT 2928
DB 2399 TGTGGGCGCTGATGCGAGGAGGTGAGGGGTGTCACCTGGGGGAGCCACAGCTGT 2458
QY 2929 GGTGACAAAGGCTTTCCTGTAGTTTATGCCCTTTGTCACCTACCTACCCCAACCACTG 2988
DB 2459 -----ACAAAGGCTTTGCTGTAGCTCTGCCCTTTGCCACTTATCCCAACCACTG 2514
QY 2989 GCTCGGCATGAAAGACTCAGTGTGAGAAACCTTCCGCTGTGCGCTGTGCTGTTCATAC 3048
DB 2515 GCGCGGCACATGAAAGACACACAGTGTGAGAGCCCTTCGCTGCCCGCTGTCTCTAT 2574
QY 3049 GCCTCTGCTCATCTGATTAACCTGAAACGSCACAGCGGTCCACACAGAGAGAAAGCC 3108
DB 2575 GCCTCTGCTCATCTGATTAACCTGAAACGSCACAGCGGTCCATACAGAGAGAGAGCC 2634
QY 3109 TACAAGTGCCTCTGTGCTGATGCTGTGGAACTGCGCAACCTTGAAGCTCATGCT 3168
DB 2635 TACAAGTGCCTCTGTGCTTATGCTGTGGCAATCTGCGCAACCTTGAAGCTCATGCT 2694
QY 3169 CGCATCCACTCTGTGACAAACCTTTTGGTGTAGCTTTGCAACTACAGCTGCAACACAG 3228
DB 2695 CGCATCCACTCTGTGACAAACCTTTTGGTGTAGCTTTGCAACTACAGCTGCAACACAG 2754
QY 3229 AGTATGAACCTCAACGCTCATATGCTGCGACACACAGGGGAGAGCCCTTCCGCTGTGC 3288
DB 2755 AGCATGAACCTCAACGCTCATATGCTGCGACACACAGGGGAGAGCCCTTCCGCTGTGC 2814
QY 3289 ACCTGGCCTTATCCACAGCGCTGAGCACTTGAAGCTTATCAGAGGTGATGCTGCG 3348
DB 2815 ACCTGGCCTTATCCACAGCGCTGAGCACTTGAAGCTTATCAGAGGTGATGCTGCG 2874
QY 3349 CATGGTGGAGCAGAGGGCTGTGCTCTTGGCTTGTAGGGCTGGGCGCCACCTCATAGC 3408
DB 2875 CACGTGGGCGAGAGGGCTGTGCTCTCTGCTTGTAGGGCTGGGCGCCACCTCATAGC 2934
QY 3409 CCACCTCTGTGTTTGAACACTCGGGGTGAGAGCCCTTGGTGTCTACTGTAGAGGGCT 3468
DB 2935 CCACCTCTGTGTTTGAACACTCGGGGTGAGAGCCCTTGGTGTCTACTGTAGAGGGCT 2994
QY 3469 CTTTATTCAGACTCACCTTGAACCTAAGTGTCTTTTACCTGGGGCTCTAGGAATTAGC 3528
DB 2995 GTCCACACAGACTCTCTGAC-----TAGG 3021
QY 3529 CCTATGCTCTGCTGATTTTATACAAATGAACCTAGAAACCACTTTCCTTTCTCCCCCGC 3588

Db 3022 TCCTTCTCCCATGTTTATATACAGACGGACCAAGACACCTT-----TTTCTCCCCCGC 3077
QY 3589 TGGTCAGGGGCTCCACACAGACTTACCTAGGCACTATATGGCCAGGCTGATCCCATGG 3648
Db 3078 TGGCCAGGGGCTCCACACAGACTTACCTAGGCACTATATGGCCAGGCTGATCCCATGG 3137
QY 3649 TCAGGGGGG-CCATATAGACACAGGGGA-CTTGTCTTAGCTTCACTACCAAGATGAGCTAAGT 3706
Db 3138 CGGGGGGGGCTATATGACACAGGGGACCTTGCCTTGACTGAGGCACTTCCAGGCTCAGT 3197
QY 3707 GATTAGGGCTTGGATTCACCGGCACTGCTCCAGAGGCTATGGAATGAAGTGGTGGAG 3766
Db 3198 GAGAGGGGCTTGTATTCACCTCCACTGCCCCAGGGGCTGTGGACAACCGGCTGGGGG 3257
QY 3767 -CTGCCCGAGCTTTTAC-----TGTTTTAACTTATTTCACTGCTTTATATAAAGGAACA 3821
Db 3258 ACTGCCCGAGCTTCCACCTGCTTTTATTAATTAATTAATTAAGGAACA 3317
QY 3822 CTAACAGA 3829
Db 3318 CTAACAAA 3325

RESULT 12
ABL97831
ID ABL97831 standard; DNA; 9266 BP.
XX
AC ABL97831;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2483.
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200156317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US000329.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 26-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232376P.
PR 14-SEP-2000; 2000US-0232386P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241231P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246603P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249298P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483232/52.
 XX
 DR Nucleic acids encoding 973 human testicular antigen polypeptides, useful
 PT for preventing, diagnosing and/or treating testicular cancer.
 XX
 PS Disclosure; SEQ ID NO 2483; 766pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a DNA encoding a
 CC protein fragment of the invention
 XX
 SQ Sequence 9266 BP; 2324 A; 2431 C; 2502 G; 2009 T; 0 U; 0 Other;
 Query Match 31.1%; Score 1195; DB 4; Length 9266;
 Best Local Similarity 81.3%; Pred. No. 2.9e-294;
 Matches 1568; Conservative 0; Mismatches 216; Indels 144; Gaps 10;
 QY 2003 AAGAGGAGAGAGAGAGATGAAGTGACGGGCCAACACACAGCTCATGGCTTTGAGA 2062
 DB 1441 AAGTGAAAGGGAGAGTTGGAGCGCTGATCCTACCTATGCTGATGTCCTTATGT 1500
 QY 2063 GAGACTCTGAGGAGACTCTCAGGGGCCAGACTTGACTTCCCTATGGCTGAGTGACG 2122
 DB 1501 CTATTTACAGGAGACTCTCTGGGGGCCAGGCTTCCCTATGGCTGAGCGACG 1560
 QY 2123 ACGAGTCTGGGGCGGCCCGCACCTAAGTGGGAGAGTGAAGTTGAGGAACACAGCCAGGG 2182

DB 1561 ATGAGTCTGGGGCGGCCGGGCACTAAGTGGGAGAGTGAAGTTGAGGAGCCAGCCAGG 1620
 QY 2183 GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCTGTGTGAGTGTGTGGGGGCCGA 2242
 DB 1621 GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCTGTGTGAGTGTGTGGGGGCCGA 1680
 QY 2243 CAGGTGAGGGGCGGTGTTGTGGGGCAGGAGGCGGGGTGGGGGGCCCCCGCTGCCCCCCAC 2302
 DB 1681 CAGGTGAGGGGCGGTGTTGTGGGGCAGGAGGCGGGGTGGGGGG-CCCTGTGTCCTCCAC 1739
 QY 2303 GGTACTGTACTATGCGCGCTTGTGGCTTTCGTGTCCCACTACTCGAGCCACCTGAAGC 2362
 DB 1740 GGCTACTGTACTATGCGCGCTTGTGGCTTTCGTGTCCCACTACTCGAGCCACCTGAAGC 1799
 QY 2363 GGCACATGACAGACACACAGCGGGGAGAGCCTGTCCGTGTGGCGCTGCCCGCTACGONT 2422
 DB 1800 GGCACATGACAGACACACAGCGGGGAGAGCCTGTCCGTGTGGCGCTGCCCGCTACGONT 1859
 QY 2423 CAGCCAGTTCGTAACTGTGA GCGGACATA CCGCACCCATACTGGCGGAGAGCCCTACC 2482
 DB 1860 CAGCCAGTTCGTAACTGTGA GCGGACATA CCGCACCCATACTGGCGGAGAGCCCTACC 1919
 QY 2483 GTTGTCCCGCAGTCCCTTTCCTGTCAGCAGCCTGCGGCAACCTGAGGGGCGCATCAGCGCA 2542
 DB 1920 GCTGTCCCGCAGTCCCTTTCCTGTCAGCAGCCTGCGGCAACCTGAGGGGCGCATCAGCGTA 1979
 QY 2543 CCACACAGGGGCTCCCACTCTCTCCCTGCCCAACCTGTGGCTTTCGATGCTGTGTCCAC 2602
 DB 1980 CCCACGAGGGCCCCCCTCTCTCTCTGCCGACCTGTGGCTTTCGCTGTACTCCAC 2039
 QY 2603 GACCAACCCGGGCTCCCACTCTCTCTCTGCCGACCTGTGGCTTTCGCTGTACTCCAC 2651
 DB 2040 GACCAGCCGGGCTCCCACTCTCTCTCTGCCGACCTGTGGCTTTCGCTGTACTCCAC 2099
 QY 2652 ----- 2651
 DB 2100 TAAGACACACAGGACCAAGATCTTGGGACATGGGTGGCTGACCTAGGAATGCTTGG 2159
 QY 2652 -----ACGATCAGAAATGGCTGATCTCTGCCAGCTTGA 2688
 DB 2160 ATTGGATTATAGCCAGGTCTTTGTCCCCACAGATGCTCTCTCTCCAGATTGAGC 2219
 QY 2689 CTTTATGTGCGACCAAGGTGTCCTTCCAGACTGTGGGCGAGTGGGGGTGAA 2748
 DB 2220 CTCATGTGCCACAGGT-GTGCCAGTTTCTTCCAGACTGTGGGCGAGTGGGGGTGAA 2278
 QY 2749 GGGGAGAGCTGTGTGGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCTTTGGACC 2808
 DB 2279 GGGGAGAGCTGTGTGGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCTTTGGACC 2338
 QY 2809 TGCCTGGGCTGTGACAGGAACTGGAGGAGGTGAGGGCAGCAGGCTGGGAGCTGCCATG 2868
 DB 2339 TGCCTGGGCTGTGACAGGAACTGGAGGAGGTGAGGGTAGTTCGGCTGGGAGCTGCCATG 2398
 QY 2869 TGTGGGCTGTGATGCGAGAGAGGCTGGAGGGTTGCCACTGGGGGAGCCCGAGGGCCCT 2928
 DB 2399 TGTGGGCTGTGATGCGAGAGAGGCTGGAGGGTTGCCACTGGGGGAGCCCGAGGGCTGT 2458
 QY 2929 GGTGACAAAGGCTTTCCTGTAGTTTATGCCCTTTGCCACTCACTTACCCCAACACCTG 2988
 DB 2459 -----ACAAAGGCTTTCCTGTAGCTTCTGCCCTTTGCCACTCACTTATCCCAACACCTG 2514
 QY 2989 GCTCGGCACATGAAGACTCACAGTGTGTGAGAAACCTTCCGCTGTGCCCGCTGTCCATAC 3048
 DB 2515 GCCCGGCACATGAAGACACACAGTGTGTGAGAAAGCCCTTCCGCTGTGCCCGCTGTCCAT 2574
 QY 3049 GCCTCTGTCTCTGTGATTAACCTGAAACCGGACCGGCTCCACACAGGAGAAAGCCC 3108
 DB 2575 GCCTCTGTCTCTGTGATTAACCTGAAACCGGACCGGCTCCATACAGGAGAAAGCCC 2634
 QY 3109 TACAAGTGTCCCTCTGTCTGTGATGCTGTGGCAACCTGGCCAACTCTCAAGGCTCATGT 3168
 DB 2635 TACAAGTGTCCCTCTGTCTGTGATGCTGTGGCAATCTGGCCAACTCTCAAGGCTCATGT 2694

QY 2652 ----- 2651
Db 7167 TAAGACACACAGGACCAAGATCTTTGGGACATGGGTGGCTGACCTAGGAATGCTGG 7108
QY 2652 -----ACGATCAGAAAAATCGCTGATCTCTGCGCAGACTTTGAGT 2688
Db 7107 ATTGGATTATAGCCAGGCTTTTGTCTCCACAGATGCTCTGCTCTTCCAGATTGGAGC 7048
QY 2689 CTTTCATGTGCGACACAGTGTGCGAGTTTCTGCGCAGACTGTGCGGAGCTGCGGGGTGAA 2748
Db 7047 CTCATGTGCGACACAGT -GTGCGCAGTTTCTGCGCAGACTGTGCGGAGCTGCGGGGTGAA 6989
QY 2749 GGGGAGAGCTTGTGTGAATCGGATCCGAACCACTGCGCAGAGTACTGTTCCTCTGGACC 2808
Db 6988 GGGGAGGGGCTCTGCGGAGCTGGATCAGAACCACTGCGCAGAGTACTGTTCCTCTGGACC 6929
QY 2809 TGGCGGGCTGTGTGACAGGAACTGGAGAGGGTGGGCGAGGCTGGGAGCTGGCATG 2868
Db 6928 TGGCGGGCTGTGTGACAGGAGCTGGAGAGGGTGGGCGAGGCTGGGAGCTGGCATG 6869
QY 2869 TGTGGGCGCTGATGCGAGGAGAGGCTGGAGGGTGGGCGAGGCTGGGAGCTGGCATG 2928
Db 6868 TGTGGGCGCTGATGCGAGGAGAGGCTGGAGGGTGGGCGAGGCTGGGAGCTGGCATG 6809
QY 2929 GGTGACAAAGGCTTTGCTGTGATTTATGCCCCCTTTGCCCCCTCACTACCCCAACCACTG 2988
Db 6808 ----ACAAAGGCTTTGCTGTGATGCTCTGCCCCCTTTGCCCCCTCACTATCCCAACCACTG 6753
QY 2989 GCTCGGCACATGAGACATCAGTGTGGAGAACTTTCGCTGTGCGGCTGTCATATC 3048
Db 6752 GCCCGGCACATGAGACACACAGTGTGGAGAGGCTTTCGCTGTGCGGCTGTCATAT 6693
QY 3049 GCCTCTGCTCATCTGGATTAACCTGAAACGGCAGCAGCGCTCCACACAGAGAGAAAGCCC 3108
Db 6692 GCCTCTGCTCATCTGGATTAACCTGAAACGGCAGCAGCGCTCCACACAGAGAGAAAGCCC 6633
QY 3109 TACAAGTGGCCCCCTGTGCTGTATGCTGTGGGAACTTGGGCAACCTCAAGCGTCATGCT 3168
Db 6632 TACAAGTGGCCCCCTGTGCTGTATGCTGTGGGAACTTGGGCAACCTCAAGCGTCATGCT 6573
QY 3169 CGCATCCACTCTGTGTGACAAACCTTTTCGGGTGATGCTTTGCAACTACAGCTGCAACAG 3228
Db 6572 CGCATCCACTCTGTGTGACAAACCTTTTCGGGTGATGCTTTGCAACTACAGCTGCAACAG 6513
QY 3229 AGTATGAACCTCAACAGCTATGCTGTGCGACACAGGCGAGAGCCCTCCGCTGTGCG 3288
Db 6512 AGCATGAACCTCAACAGCTATGCTGTGCGACACAGGCGAGAGCCCTCCGCTGTGCG 6453
QY 3289 ACCTGGGCTATACACAGGCGCACTGGGACAACTACAGCGCTCATCAGAGGTCATGGC 3348
Db 6452 ACCTGGGCTATACACAGGCGCACTGGGACAACTACAGCGCTCATCAGAGGTCATGGC 6393
QY 3349 CATGTGTGAGCAGGAGGCGCTGCTCTCTGCCCCCTGAGGCTGGGCCCCACCTCATAGC 3408
Db 6392 CACGTGGGCGAGGAGGCGCTGCTCTCTGCCCCCTGAGGCTGGGCCCCACCTCATAGC 6333
QY 3409 CCACCTCTGTTTGTAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGTGTAGCAGGCT 3468
Db 6332 CCACCTCTGTTTGTAGCTCTGGGGCCACAGCCCTGGGAGCTGTGTGGCAGCGGCT 6273
QY 3469 CTTTCATTGAGACTCACCTTGAACCTAACTAGGTTCTTTTACCTGGGCTCTAGGAATTAGC 3528
Db 6272 GTCCACACAGACTCATCTGTAAC-----TAGG 6246
QY 3529 CCTATGCTCTCTGCAATTTTATACAAATGAATAGAAACAGCTTTCTCTCTGCCCCGC 3588
Db 6245 TCT 6190
QY 3589 TGGTCAGGGGCTCCACACAGACTTAACCTAGGCACTATATGACCAAGCTGTAATCCCATGG 3648
Db 6189 TGGCCAGGGGCTCCACACAGACTTAACCTAGGCACTATATGACCAAGCTGTAATCCCATGG 6130

QY 3649 TCAGGGGG-CCATATAGACACAGGGGA-CTTGTCTTAGCTACGTACCAAGATGAGCTAAGT 3706
Db 6129 CGGGGGGGCCCATATGACACAGGGGACCTTGCCTTGAAGGACCTTCAAGGCTCAGT 6070
QY 3707 GATTAGGGCTTGGATTACCGCCCACTGCTCCAGAGGCTATGGATGAACCTGTTGGGAG 3766
Db 6069 GAGAAAGGGCCCTGTATTACCTCCACTGCCCCAGGGGCTGTGGACAAACCGCTGGGGG 6010
QY 3767 -CTGCCAGGCTTTTAC-----TGTTTTAACTTATTTCAGTCTTTTATATAAGGAAACA 3821
Db 6009 ACTGCCAGGCTCCCACTGTTTATTAACTTATTTCAGTCTTTTATATAAGGAAACA 5950
QY 3822 CTAACAGA 3829
Db 5949 CTAACAAA 5942

RESULT 14
ADA53558
ID ADA53558 standard; cDNA; 1456 BP.
XX
AC ADA53558;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human coding sequence, SEQ ID 1126.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX Homo sapiens.
XX
XX BP1293569-A2
XX 19-MAR-2003.
XX
XX 21-MAR-2002; 2002EP-00006586.
XX
XX 14-SEP-2001; 2001JP-00328381.
XX 24-JAN-2002; 2002US-0350435P.
XX
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Naganari K, Masuho Y;
XX
XX WPI: 2003-395539/38.
XX P-PSDB; ADA55197.
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
XX and/or membrane proteins, useful for developing medicines for diseases in
XX which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 1; SEQ ID NO 1126; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
XX proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
XX ADA54071). The coding sequences are useful in the gene therapy of
XX diseases caused by abnormalities of the proteins, e.g. cancer,
XX inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 1456 BP; 287 A; 495 C; 406 G; 268 T; 0 U; 0 Other;

Query Match 29.0%; Score 1115.2; DB 7; Length 1456;
Best Local Similarity 87.4%; Pred. No. 2.8e-274; Indels 44; Gaps 6;
Matches 1305; Conservative 0; Mismatches 144;

QY 2341 CACTACTCGAGCCACCTGAGCGGCACATGACACACACAGCGGAGAGCGGTTCCGC 2400
Db 1 CACTACTCGAGCCACCTGAGCGGCACATGACACACACAGCGGAGAGCGGTTCCGC 60

Qy 2401 TGTGGCGCTGCGCCATACGTCNTCAGCCCACTTCGTCACCTGACGCGACATACCGCACC 2460
Db 61 TGTGGCGCTGCGCCATACGTCNTCAGCCCACTTCGTCACCTGACGCGACATACCGCACC 120
Qy 2461 CATACTGGGAGAGCCCTACCGTTGTCCCACTGCGCCCTTTGCGTGCAGAGCCCTGGC 2520
Db 121 CACACTGGGAGAGCCCTACCGTTGTCCCACTGCGCCCTTTGCGTGCAGAGCCCTGGC 180
Qy 2521 AACCTGAGCGGCGATCAGCGGCAACCCACAGAGGCGCTCCCACTCCCTCCCTGCCCAACCTGT 2580
Db 181 AACCTGAGCGGCGATCAGCGTACCCACAGCGGCGCCCACTCCCTGCCCGACCTGT 240
Qy 2581 GGTCTTCGATGCTGCTGCTCAAGCAACCCGCGCTCCCACTCCCACTCCCACTGAGAGGCG 2640
Db 241 GGTCTTCGCTGCTGCTCAAGCAACCCGCGCTCCCACTCCCACTCCCACTGAGAGGCG 300
Qy 2641 ACAATGCCCGGCGATCAGAGGCGCTGATCCCTGCGCAGACTTCAGTCTTCATGTGCCA 2700
Db 301 GCGGTGCGCGGCGACTGAGATGCTGCTCTTCCAGATTTGAGCCTCCATGTGCCA 360
Qy 2701 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Db 361 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 2761 TGTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 421 TGTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 2821 GGCAGAGAACTGGAGAGGCTGAGGCGAGCGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTG 2880
Db 481 GGCAGAGAACTGGAGAGGCTGAGGCGAGCGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTG 540
Qy 2881 ATCGGAGGAGCTGGAGGCTGAGGCGAGCGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTG 2940
Db 541 ATCGGAGGAGCTGGAGGCTGAGGCGAGCGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTG 600
Qy 2941 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 601 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 3001 AAGACTCAGTGTGAGAAACCTTCCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
Db 661 AAGACTCAGTGTGAGAAACCTTCCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 3061 CTGGATACTGAAACCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 721 CTGGATACTGAAACCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 3121 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Db 781 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 3181 GGTGACAAACCTTTTGGTGTAGCTTGGCACTGACAGCTGCAACAGAGATGAACTTC 3240
Db 841 GGTGACAAACCTTTTGGTGTAGCTTGGCACTGACAGCTGCAACAGAGATGAACTTC 900
Qy 3241 AAAGCTCATATGCTGCAACACAGGCGAGAGCCCTTCCGCTGCTGCTGCTGCTGCTGCTGCT 3300
Db 901 AAAGCTCATATGCTGCAACACAGGCGAGAGCCCTTCCGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 3301 ACCACAGGCGCTGGGCAACTACAGCTCATCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Db 961 ACCACAGGCGCTGGGCAACTACAGCTCATCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 3361 GGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Db 1021 GGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 3421 TTGAGCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db 1081 TTGAGCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

Qy 3481 TCACCTGAACTAAGTGGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCCT 3540
Db 1141 TCATCTGAACTAAGTGGTCTTTCTTC-----CCC 1167
Qy 3541 GCATTTTATCAAAATGAATAGAAACACCTTTTCCCTTTCTCCCCCGCTGGTCAAGGGCT 3600
Db 1168 ATGTTTATACAGCGGACCGAGCCACCTT-----TTTCTCCCCGCTGGCCAGGGCT 1223
Qy 3601 CCACACAGATAAAGTGGTCTATATGAGCAGAGCTGATCCATGCTGCTGAGGGG-CCA 3659
Db 1224 CCACACAGATAAAGTGGTCTATATGAGCAGAGCTGATCCATGCTGCTGAGGGGCCCCA 1283
Qy 3660 TATAGACACAGGGA-CTTGTCTTAGCTCAGCTACAGATGAGCTAAGTGAATTAGGGCCCTT 3718
Db 1284 TATGACACAGGAGACCTTGTCTTGAAGTGGTCTGAGCTCAGTGAGAGGGCCCT 1343
Qy 3719 GGAATCACCCGCTGCTCCAGAGGCTATGAGTGAAGTGGTGGAG-CTGCCAGGCT 3777
Db 1344 GTATTACCTCCACTGCGCCCGAGGGCTGTGGACAAACCGCTGGGGGAGTGGCCAGCT 1403
Qy 3778 TTAC-----TGTTTACTTATTTTCACTGCTTTTATATTAAGGAAACACTAAC 3826
Db 1404 CCCACCTGTTTATTTAACTTATTTTCACTGCTTTTATATTAAGGAAACACTAAC 1456

RESULT 15
AAI58156
ID AAI58156 standard; cDNA; 1558 BP.
XX AAI58156;
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 359.
XX Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou F, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM39000.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Claim 1; SEQ ID NO 359; 10078pp; English.

XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

Sequence 1558 BP; 331 A; 456 C; 494 G; 277 T; 0 U; 0 Other;

Query Match	21.2%	Score 815.2	DB 4	Length 1558
Best Local Similarity	85.7%	Pred. No. 1.4e-197		
Matches 1004: Conservative		0	Mismatches 123	Indels 45
Matches 1004: Gaps				7

2674	QY	CTGCCAGCTTGAGTCTTTCATGTGCCACCAAGTGGTGCCAGATTTCTGTGCCAGACTGTGGG	2733
425	Db	CCGACAGATTGAGCCCTCCATGTGTGCCACCAAGTGGTGCCAGATTTCTGTGCCAGACTGTGGG	484
2734	QY	CAGCTGGGGTGGAAGGGGAGAGCTTGTGTGGAACATGGAATCCGAAACAATCTGCCAGAGCTA	2793
485	Db	CAGCTGGGGGTGAAGGGGAGGGCCCTCTGCGGACTGGATCAGAAACAATGCCAGAGCTG	544
2794	QY	CTGTTCCTCTTGGACCTCCGGGGCTGTGGACAGGAACATGGAGAGGGTGAAGGACAGAG	2853
545	Db	CTATTCCCTTGGACCTCCGGGGCTGTGGACAAGAGCTGGAGGAGGTGAGGTAATCTCGG	604
2854	QY	CTGGGAGCTGCCATGTGTGGGCGCTCATGCGGAGGAGAGCTGGAGGGGTGCCACTGGG	2913
605	Db	CTGGGAGCTGCCATGTGTGGGCGCTCATGCGGAGGAGGCTGGAGGGGGTGCACATGGG	664
2914	QY	GGACCCAGAGGCCCTGTGTGACAAAGGCTTTGCCCTGTAGTTATTGCCCTTTGCCACTCAC	2973
665	Db	GGGCCCCAGGGCCCCAGTGACAAAGGCTTTGCCCTGTAGCCCTTGCCCTTTGCCACTCAC	724
2974	QY	TACCCCCAACCACTGGGTCCGGACATGAAGACTCACAGTGGTGAGAAACCCCTTCCGCTGT	3033
725	Db	TATCCCCAACCACTGGGCCGGCACATGAAGACACACAGTGGTGAGAGCCCTTCCGCTGC	784
3034	QY	GCCCGCTGTCCATAGCCCTGTGCTCATCTGATAAATCGAAAGCGGACACAGCCGCTCCAC	3093
785	Db	GCCCGCTGTCTTATGCCCTCTGCTCATCTGTGATAAATCGGACACAGCCGCTCCAT	844
3094	QY	ACAGGAGAAAGCCCTCAAGTGCCGCCCTCTGTCCGTATGCCCTGTGGCAACCTGGCCAAC	3153
845	Db	ACAGGAGAGGCCCTCAAGTGC GCCCTCTGCTGCGCAATCTTGCCCAATCTGGCCAAC	904
3154	QY	CTAAGGGTCAATGTGGATGCCATCTGTGGTGACAAACCTTTTCGGTGTAGCCTTTGCCAAC	3213
905	Db	CTAAGGGTCAATGTGGATGCCATCTGTGGTGACAAACCTTTTCGGTGTAGCCTTTGCCAAC	964
3214	QY	TACAGCTCAACACAGAGTATGAACCTCAAGCGTCAATGTGTGCACACACGGCGAGAAG	3273
965	Db	TACAGCTCAACACAGAGCATGAACCTCAAGCTCAATGTGTGCAGCACACAGGCGAGAAG	1024
3274	QY	CCCTTCGCTGTGCCAATCTGCGCTATACCAAGGCCACTGGGACAACTCAAGAGCTCAT	3333
1025	Db	-CCTTCGCTGTGCCAATCTGCGCTATACCAAGGCCACTGGGACAACTCAAGAGCTCAT	1083
3334	QY	CAGAGGTGCATGCCATGTGTGGACGAGGGGCTGTGCTCTCTGCGCCCTGAGGGCTGG	3393
1084	Db	CAGAGGTGCATGGCCACGGTGGGGCAGAGGGGCTGTGCTCTCTGCGCTGTAGGGCTGG	1143
3394	QY	GCCCCACCTCATAGCCCAACCTCTGTTTGTAGCACTCGGGGTCAGAGAGCCCTGGGCTCT	3453

Search completed: May 13, 2004, 23:07:59
Job time : 1399 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 13:25:40 ; Search time 14620 Seconds
(without alignments)
11404.969 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 ggaggttgagatttattc.....gaaaaaaaaaaaaaaaaa 3847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pl.*
25: em_ro.*
26: em_scs.*
27: em_un.*
28: em_un.*
29: em_un.*
30: em_hlg_hum.*
31: em_hlg_inv.*
32: em_hlg_Other.*
33: em_hlg_mus.*
34: em_hlg_pln.*
35: em_hlg_rod.*
36: em_hlg_mam.*
37: em_hlg_vrt.*
38: em_sy.*
39: em_hlg_hum.*
40: em_hlg_mus.*
41: em_hlg_Other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3836	99.7	3847	6	BD234712	BD234712 CNRE bind
2	2815.4	73.2	19296	2	AC114619	AC114619 Mus muscu
3	2628.4	68.3	2651	6	BD234722	BD234722 CNRE bind
4	2288	59.5	2289	6	BD234713	BD234713 CNRE bind
5	2092	54.4	274060	2	AC095184	AC095184 Rattus no
6	1988.4	51.7	2336	6	BD234721	BD234721 CNRE bind
7	1918.8	49.9	2220	10	BC058976	BC058976 Mus muscu
8	1518	39.5	2013	6	AX834458	AX834458 Sequence
9	1518	39.5	2013	9	AK096933	AK096933 Homo sapi
10	1505.8	39.1	2159	9	BC052282	BC052282 Homo sapi
11	1498	38.9	1975	6	AX274856	AX274856 Sequence
12	1367.6	35.5	1996	6	BD242880	BD242880 Secretd
13	1255.2	32.6	195880	9	AC074117	AC074117 Homo sapi
14	1115.4	29.0	1461	9	BC040650	BC040650 Homo sapi
15	1115.2	29.0	1456	6	AX714442	AX714442 Sequence
16	1115.2	29.0	1456	9	AK056765	AK056765 Homo sapi
17	960.4	25.0	112679	2	AC142219	AC142219 Rattus no
18	960.4	25.0	300078	2	AC109067	AC109067 Rattus no
19	835	21.7	1322	9	HSN805126	AL833946 Homo sapi
20	815.2	21.2	1558	6	AR338541	AR338541 Sequence
21	615.8	16.0	1824	6	AR083379	AR083379 Sequence
22	615.8	16.0	2069	10	MMU42383	U42383 Mus musculu
23	615.8	16.0	2181	10	BC009004	BC009004 Mus muscu
24	612.8	15.9	49743	2	AC114541	AC114541 Mus muscu
25	576.8	15.0	581	6	BD234719	BD234719 CNRE bind
26	559.6	14.5	49743	2	AC114541	AC114541 Mus muscu
27	528.4	13.7	851	6	BD082405	BD082405 87 human
28	526.8	13.7	2182	10	BC062083	BC062083 Rattus no
29	512.2	13.3	69843	2	AC025903	AC025903 Homo sapi
30	498	12.9	509	6	BD234714	BD234714 CNRE bind
31	485.8	12.6	500	6	BD234715	BD234715 CNRE bind
32	472.8	12.3	487	6	BD234716	BD234716 CNRE bind
33	450.2	11.7	586	6	BD234720	BD234720 CNRE bind
34	441.8	11.5	446	6	BD234717	BD234717 CNRE bind
35	399.2	10.4	496	6	AX400960	AX400960 Sequence
36	387.2	10.1	721	9	BC014386	BC014386 Homo sapi
37	387.2	10.1	1425	9	BC007361	BC007361 Homo sapi
38	385.6	10.0	1435	6	AX615059	AX615059 Sequence
39	385.6	10.0	2160	9	BC022061	BC022061 Homo sapi
40	385.6	10.0	2167	9	BC000057	BC000057 Homo sapi
41	385.6	10.0	2169	9	AK127593	AK127593 Homo sapi
42	383.6	10.0	706	6	BD021657	BD021657 Novel gen
43	383.6	10.0	706	6	BD101595	BD101595 Novel gen
44	380	9.9	809	6	AX053343	AX053343 Sequence
45	346.8	9.0	441	6	BD234718	BD234718 CNRE bind

ALIGNMENTS

RESULT 1
BD234712
LOCUS
BD234712
DEFINITION
BD234712
ACCESSION
BD234712
VERSION
BD234712.1
KEYWORDS
JP 2002516664-A/6.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3847)
AUTHORS
Chen, Y.B., Horiuchi, M., Dzu, V.J. and Tamura, K.
TITLE
CNRE binding factors and uses thereof
JOURNAL
Patent: JP 2002516664-A 6 11-JUN-2002;

BD234712 3847 bp DNA linear PAT 17-JUL-2003
CNRE binding factors and uses thereof.

[illegible]

Db 1741 GAGGGGACCTCTCAGGGTTAGGGCGGATGATCTGGGATCTTCCTTACAGAGTA 1800
Qy 1801 TTAAGGAACCTGAGGGTCATCAGGTACGGAAGTGCAGTTCACAACAGCTGGCTCCTTG 1860
Db 1801 TTAAGGAACCTGAGGGTCATCAGGTACGGAAGTGCAGTTCACAACAGCTGGCTCCTTG 1860
Qy 1861 GTTCGGATTATGGTACTGCTTGGGAGGAGATTCCACAAGACACCTCCCTCTTTAGTG 1920
Db 1861 GTTCGGATTATGGTACTGCTTGGGAGGAGATTCCACAAGACACCTCCCTCTTTAGTG 1920
Qy 1921 GATACCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGGAGACGATTTGCTA 1980
Db 1921 GATACCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGGAGACGATTTGCTA 1980
Qy 1981 CTAGGCCAAGATCTGAGATTGGAAGAGGAGAGGAGAGATGAAGGTACGGCCACAAC 2040
Db 1981 CTAGGCCAAGATCTGAGATTGGAAGAGGAGAGGAGAGATGAAGGTACGGCCACAAC 2040
Qy 2041 GACACGCTCATGGGCTTTGAGAGAGACTCTCTGAAGGAGACTCTCAGGGGCCAGACCTGGA 2100
Db 2041 GACACGCTCATGGGCTTTGAGAGAGACTCTCTGAAGGAGACTCTCAGGGGCCAGACCTGGA 2100
Qy 2101 CTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGCGGCGCGCACTAAAGTGGGAGGT 2160
Db 2101 CTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGCGGCGCGCACTAAAGTGGGAGGT 2160
Qy 2161 GAAGTTGAGGAAACAGCAGGGGTCAGGGGAGGGCCAGGGGTGAGAGGCCAGGCCCAGGC 2220
Db 2161 GAAGTTGAGGAAACAGCAGGGGTCAGGGGAGGGCCAGGGGTGAGAGGCCAGGCCCAGGC 2220
Qy 2221 TGTGAGCTGTGGGGGCCGCAAGGTGAGGGGCCCTGTGTGGGGCAGAGAGGGGGGT 2280
Db 2221 TGTGAGCTGTGGGGGCCGCAAGGTGAGGGGCCCTGTGTGGGGCAGAGAGGGGGGT 2280
Qy 2281 GGGGGGCCCGCTGCCCGACAGGTTACTGTACTCATGCCGCTGCGCTTCTGTTGCC 2340
Db 2281 GGGGGGCCCGCTGCCCGACAGGTTACTGTACTCATGCCGCTGCGCTTCTGTTGCC 2340
Qy 2341 CACTACTGAGGCACTGGAAGCGGCAATGACAGCAACAGAGGGGAGAGCGTTTCGC 2400
Db 2341 CACTACTGAGGCACTGGAAGCGGCAATGACAGCAACAGAGGGGAGAGCGTTTCGC 2400
Qy 2401 TGTGGCGCTGCCATACCGNTCAGCCGAGTTCGTCAACTGACGAGCATACCCGACC 2460
Db 2401 TGTGGCGCTGCCATACCGNTCAGCCGAGTTCGTCAACTGACGAGCATACCCGACC 2460
Qy 2461 CATACTGGCAGAGCCCTACCGTTGTGCCCATGCCCCCTTTGCTGACAGCCTGGGC 2520
Db 2461 CATACTGGCAGAGCCCTACCGTTGTGCCCATGCCCCCTTTGCTGACAGCCTGGGC 2520
Qy 2521 AACCTGAGCGGCGATCAGCGCACCCACAGGGGCTCCCACTCTCCTCGCCCACTGT 2580
Db 2521 AACCTGAGCGGCGATCAGCGCACCCACAGGGGCTCCCACTCTCCTCGCCCACTGT 2580
Qy 2581 GGCTTTGATGTGTCTCCACGACCAACCCGCGCTCCCACTCCACAGAGAGAGGGG 2640
Db 2581 GGCTTTGATGTGTCTCCACGACCAACCCGCGCTCCCACTCCCAAGAGAGAGGGG 2640
Qy 2641 ACAATGCCCGCAGATCAGAAAATGGCTGATCCTGCGAGACTTGATGTCTTATGTGCCA 2700
Db 2641 ACAATGCCCGCAGATCAGAAAATGGCTGATCCTGCGAGACTTGATGTCTTATGTGCCA 2700
Qy 2701 CCAGTGGTGCAGTTTCTCGCAGACTGTGGGCGAGCTGGGGGTGAAGGGGAGAGCTTG 2760
Db 2701 CCAGTGGTGCAGTTTCTCGCAGACTGTGGGCGAGCTGGGGGTGAAGGGGAGAGCTTG 2760
Qy 2761 TGTGGAACTCGATCCGAAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCCGGGCTGT 2820
Db 2761 TGTGGAACTCGATCCGAAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCCGGGCTGT 2820
Qy 2821 GGACAGAACTGGAGGGGTGAGGGCAGCAGCTGGAGCTGCCATGTGTGGGCGCTGC 2880

Db 2821 GGACAGAACTGGAGGGGTGAGGGCAGCAGGCTGGAGCTGCCATGTGTGGGCGCTGC 2880
Qy 2881 ATGGAGAGAGGGTGGAGGGTTGCCACTGGGGGACCCAGGGCCCTGTGTGACAAAGGC 2940
Db 2881 ATGGAGAGAGGGTGGAGGGTTGCCACTGGGGGACCCAGGGCCCTGTGTGACAAAGGC 2940
Qy 2941 TTTGCCCTGATGTTTATGCCCCCTTTGCCACTCACTTACCCCAACACCTGGCTCGGCACATG 3000
Db 2941 TTTGCCCTGATGTTTATGCCCCCTTTGCCACTCACTTACCCCAACACCTGGCTCGGCACATG 3000
Qy 3001 AAGACTCAGTGTGAGAAACCCCTTCGGCTGTGGCCGCTGTTCATAGGCTCTGTCTCAT 3060
Db 3001 AAGACTCAGTGTGAGAAACCCCTTCGGCTGTGGCCGCTGTTCATAGGCTCTGTCTCAT 3060
Qy 3061 CTGGATTAACCTGAAACCGCACACGCGGCTCCACAGAGAGAAAGCCCTCAAGGTGCCCC 3120
Db 3061 CTGGATTAACCTGAAACCGCACACGCGGCTCCACAGAGAGAAAGCCCTCAAGGTGCCCC 3120
Qy 3121 CTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTCACTGTGCGATCCACTCT 3180
Db 3121 CTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTCACTGTGCGATCCACTCT 3180
Qy 3181 GGTGACAAACCTTTCCGCTGTAGCCTTTGCAACTACAGCTGCRACACAGGTATGAACCTC 3240
Db 3181 GGTGACAAACCTTTCCGCTGTAGCCTTTGCAACTACAGCTGCRACACAGGTATGAACCTC 3240
Qy 3241 AAAAGTCAATGCTGCGACACACGCGGGGAGAGCCCTTCGCTGTGCCACCTGCGCCTAT 3300
Db 3241 AAAAGTCAATGCTGCGACACACGCGGGGAGAGCCCTTCGCTGTGCCACCTGCGCCTAT 3300
Qy 3301 ACCACAGCCCACTGGGCAACTACAGCGTCACTAGAGGTGATGGCCATGTGTGGAGCA 3360
Db 3301 ACCACAGCCCACTGGGCAACTACAGCGTCACTAGAGGTGATGGCCATGTGTGGAGCA 3360
Qy 3361 GGAGGGCCGTGCTCTCTGCCCCCTGAGGGCTGGGCCACCTCATAGCCCCACCTCTGTT 3420
Db 3361 GGAGGGCCGTGCTCTCTGCCCCCTGAGGGCTGGGCCACCTCATAGCCCCACCTCTGTT 3420
Qy 3421 TTGAGCACTCGGGGTCCAGAGCCCTGGTGCTACTGTAGTAGAGGGCTCTTCATTAGAC 3480
Db 3421 TTGAGCACTCGGGGTCCAGAGCCCTGGTGCTACTGTAGTAGAGGGCTCTTCATTAGAC 3480
Qy 3481 TCACCTTGAACTAACTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCCCTCT 3540
Db 3481 TCACCTTGAACTAACTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCCCTCT 3540
Qy 3541 GCAITTTATACAAATGAACCTAGAAACCACTTTCCCTTTCTCCCCCGCTGGTCAGGGGCT 3600
Db 3541 GCAITTTATACAAATGAACCTAGAAACCACTTTCCCTTTCTCCCCCGCTGGTCAGGGGCT 3600
Qy 3601 CCACAGACTAACTAGGCACTATATGGACCGCTGAATCCCATGCTGTCAGGGGCCAT 3660
Db 3601 CCACAGACTAACTAGGCACTATATGGACCGCTGAATCCCATGCTGTCAGGGGCCAT 3660
Qy 3661 ATAGACAGGGGACTTGTCTTGTAGCTCAGTACCGATGAGCTAAAGTATTAGGGCCTTGG 3720
Db 3661 ATAGACAGGGGACTTGTCTTGTAGCTCAGTACCGATGAGCTAAAGTATTAGGGCCTTGG 3720
Qy 3721 ATTACCGCCACTGCTCCAGAGCTATGATGAACCTGGTTGGAGCTGCCAGCCTTTT 3780
Db 3721 ATTACCGCCACTGCTCCAGAGCTATGATGAACCTGGTTGGAGCTGCCAGCCTTTT 3780
Qy 3781 ACTGTTTTTAACTTATTTTCACTGCTTTTATATAAAGGAAACACTAACAGAAAAA 3840
Db 3781 ACTGTTTTTAACTTATTTTCACTGCTTTTATATAAAGGAAACACTAACAGAAAAA 3840
Qy 3841 AAAAAA 3847
Db 3841 AAAAAA 3847

RESULT 2
AC114619

AC114619 199296 bp DNA linear HTG 06-JUN-2002
 Mus musculus clone RP24-86123, WORKING DRAFT SEQUENCE, 15 ordered
 pieces
 AC114619
 HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 199296)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP24-86123
 Unpublished
 2 (bases 1 to 199296)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chararo,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,F., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 199296)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chararo,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
 Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,F.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
 Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
 Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S.,
 Theodore,J., Topham,K., Travers,M., Travis,N., Triggilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 6, 2002 this sequence version replaced gi:19311148.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RN/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24325
 Center clone name: 861.23
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 193147 bases at least Q40
 Consensus quality: 196185 bases at least Q30
 Consensus quality: 197247 bases at least Q20
 Insert size: 194000; agarose-1p
 Insert size: 197896; sum-of-contigs
 Quality coverage: 6.1 in Q20 bases; agarose-1p
 Quality coverage: 6.0 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * -----
 1 728: contig of 728 bp in length
 729 828: gap of 100 bp
 829 2029: contig of 1201 bp in length
 2129: gap of 100 bp
 2130 3803: contig of 1674 bp in length
 3804 3903: gap of 100 bp
 6051: contig of 2148 bp in length
 6151: gap of 100 bp
 6152 10082: contig of 3931 bp in length
 10083 10182: gap of 100 bp
 10183 18316: contig of 8134 bp in length
 18316 18416: gap of 100 bp
 18417 27114: contig of 8698 bp in length
 27114 33934: gap of 100 bp
 33934 34034: contig of 6720 bp in length
 34035 44466: contig of 10432 bp in length
 44467 44567: gap of 100 bp
 44567 59489: contig of 14923 bp in length
 59490 79430: contig of 19841 bp in length
 79431 79530: gap of 100 bp
 79531 99546: contig of 20015 bp in length
 99546 123291: contig of 23645 bp in length
 123291 123390: gap of 100 bp
 123391 151687: contig of 28297 bp in length
 151688 151788: gap of 100 bp
 151788 199296: contig of 47509 bp in length.
 ----- Location/Qualifiers
 1. 199296
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-86123"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1. 728
 /note="assembly_fragment"
 829. 2029
 /note="assembly_fragment"
 2130. 3803
 /note="assembly_fragment"
 3904. 6051
 /note="assembly_fragment"
 6152. 10082
 /note="assembly_fragment"
 10183. 18316
 /note="assembly_fragment"
 misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature

[illegible]

Qy	720	GGCCCCCACC	CCCCCCCCAC	CGCGGAGC	CTCGTGCCTT	TAAGCCTCCTT	779
Db	183065	GGCCCCCACC	CCCCCCCCAC	CGCGGAGC	CTCGTGCCTT	TAAGCCTCCTT	183124
Qy	780	ACTCTCTTGG	CGTCACTAC	CGTCTGCT	CTGTGTGT	TGGAGGAA	839
Db	183125	ACTCTCTTGG	CGTCACTAC	CGTCTGCT	CTGTGTGT	TGGAGGAA	183184
Qy	840	GGACTGTAGT	CTGATTTT	TACTCTGT	GAACTTTT	TAAGGACATCTCTTTTATT	899
Db	183185	GGACTGTAGT	CTGATTTT	TACTCTGT	GAACTTTT	TAAGGACATCTCTTTTATT	183244
Qy	900	GGGGCTCTGT	GACCCCTAG	CGCTTGC	ACCCGCTCT	CTGTGTGTACACTTTCAAGCAACA	959
Db	183245	GGGGCTCTGT	GACCCCTAG	CGCTTGC	ACCCGCTCT	CTGTGTGTACACTTTCAAGCAACA	183304
Qy	960	CTTTTTCAG	ACTAAAGCC	CAACAAAG	CTAATCGT	CTCATGTGATCTTTTACTCT	1019
Db	183305	CTTTTTCAG	ACTAAAGCC	CAACAAAG	CTAATCGT	CTCATGTGATCTTTTACTCT	183364
Qy	1020	CCTACCTTGG	CCCCATAT	GTGGCGG	CTGTTCG	TGCACTGTGTGTGCCGAGC	1079
Db	183365	CCTACCTTGG	CCCCATAT	GTGGCGG	CTGTTCG	TGCACTGTGTGTGCCGAGC	183424
Qy	1080	CTCTGGAG	AGGGGCG	AGTGAAT	GTGGAGC	AGGAGCTGGGAGTCTTAACCTCTGAGTAAGC	1139
Db	183425	CTCTGGAG	AGGGGCG	AGTGAAT	GTGGAGC	AGGAGCTGGGAGTCTTAACCTCTGAGTAAGC	183484
Qy	1140	AGCCGCTGT	GAGGCATC	CTGCGCT	GTTCGCTT	CAGCAGCACTTCTTAGTTCGTATAT	1199
Db	183485	AGCCGCTGT	GAGGCATC	CTGCGCT	GTTCGCTT	CAGCAGCACTTCTTAGTTCGTATAT	183544
Qy	1200	GGAGGGCG	CAGGAT	ACCACCTT	CTGCTGGT	TGGCTTACGCTCCAGCTTCTCAT	1259
Db	183545	GGAGGGCG	CAGGAT	ACCACCTT	CTGCTGGT	TGGCTTACGCTCCAGCTTCTCAT	183604
Qy	1260	TGTTTGGT	AGCAGG	GCATCA	CAATCTT	CCTGGTCACTGCGCTCCGGTGTG	1319
Db	183605	TGTTTGGT	AGCAGG	GCATCA	CAATCTT	CCTGGTCACTGCGCTCCGGTGTG	183664
Qy	1320	TACAGCCAT	GAACTAG	CGATCC	CGTGTCT	CGGTGCGGAAGCGGAGCGGTAC	1379
Db	183665	TACAGCCAT	GAACTAG	CGATCC	CGTGTCT	CGGTGCGGAAGCGGAGCGGTAC	183724
Qy	1380	GGAGTACCA	GTCTGG	TTCGAG	AGGGGGT	TAGGGGTCTCATGTAATGGAACGGCGCGCGG	1439
Db	183725	GGAGTACCA	GTCTGG	TTCGAG	AGGGGGT	TAGGGGTCTCATGTAATGGAACGGCGCGCGG	183784
Qy	1440	CGCGGAG	CGACCTG	AGCTGAT	TCGGGGC	CCGGG - CAGGGGTGCCCAGGGGCCCGCA	1498
Db	183785	CGCGGAG	CGACCTG	AGCTGAT	TCGGGGC	CCGGG - CAGGGGTGCCCAGGGGCCCGCA	183844
Qy	1499	CC - GTGTAT	TGGGGCG	GGTTCT	GTGATCCT	TAAGACAAGACCGCAGCGAGGGCCGAACTG	1557
Db	183845	CCGTGTAT	TGGGGCG	GGTTCT	GTGATCCT	TAAGACAAGACCGCAGCGAGGGCCGAACTG	183904
Qy	1558	GGAGGGCG	GGGCGG	AGGCTCG	GGCCGAG	CGCGTCCGGCTTGGAGCCGCTCACCAT	1617
Db	183905	GGAGGGCG	GGGCGG	AGGCTCG	GGCCGAG	CGCGTCCGGCTTGGAGCCGCTCACCAT	183964
Qy	1618	CCCCCAAG	GAAGAA	AGCCAT	CCACAG	CCCGTGAATTCGAGGGGGGTCAAGGTCAAGG	1677
Db	183965	CCCCCAAG	GAAGAA	AGCCAT	CCACAG	CCCGTGAATTCGAGGGGGGTCAAGGTCAAGG	184024
Qy	1678	GTCAAGGG	CCCTTG	AGCCGG	AGGAACAG	GGGTGGGTCAGTAGAGTGGGCTCAGGTCAGG	1737
Db	184025	GTCAAGGG	CCCTTG	AGCCGG	AGGAACAG	GGGTGGGTCAGTAGAGTGGGCTCAGGTCAGG	184084
Qy	1738	GTGAGGG	GGGACTCT	CTCAGG	GTAGGG	GTAGGGTGGGTCAGTAGAGTGGGCTCAGGTCAGG	1797
Db	184085	GTGAGGG	GGGACTCT	CTCAGG	GTAGGG	GTAGGGTGGGTCAGTAGAGTGGGCTCAGGTCAGG	184144
Qy	1798	GTATTAA	AGGAAC	CTGAGG	GTCAAT	CGGAAGTGCAGTTCAACAACACTGGCTCC	1857

Db	184145	GTATTAAGGAACCTGAGGTCATCGAGTACGGGAAGTGCAGTTCACAAACAGCTGGCTCC	184204	Db	185225	TGACACAGAGTCTGGGGGGCGCGCACA	CTAAGTTCGGAGAGTGAAGTTGAGGAACACGAC	185284
Qy	1858	TTGTTTCGGATTATGGGPACTGCTTGGGAGGAGATTCCCAAGCACCTCCCTCTTTTA	1917	Qy	2178	CAGGGGTTCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCCTGTCTCAGCTGTGTGGGG	2237	
Db	184205	TTGGTTCCGATTATGGGPACTGCTTGGGAGGAGATTCCCAAGCACCTCCCTCTTTTA	184264	Db	185285	CAGGGGTTCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCCTGTCTCAGCTGTGTGGGG	185344	
Qy	1918	GTGGATACCTGAAGATTCCCTTCGACGAAGTCTCCGGGCGCTTGGTGTGGAGAGCGATTG	1977	Qy	2238	GCCGACAGGTGAGGGGCGCTGTGTGGGCGAGGAGGGCGGGT-GGGGGGCCCCCGCTGC	2296	
Db	184265	GTGGATACCTGAAGATTCCCTTCGACGAAGTCTCCGGGCGCTTGGTGTGGAGAGCGATTG	184324	Db	185345	GCCGACAGGTGAGGGGCGCTGTGTGGGCGAGGAGGGCGGGTGGGGGGCCCCCGCTGC	185404	
Qy	1978	CTACTAGCCCAAGATCTCGAGTTTGAAGAGGAAGAGAGGATGAAGTGAAGCGGCAC	2037	Qy	2297	CCCCCGGTACTGTACTCATGCCGCTGTGGCTTTTGGTGTCCCACTACTCGAGCCACC	2356	
Db	184325	CTACTAGCCCAAGATCTCGAGTTTGAAGAGGAAGAGAGGATGAAGTGAAGCGGCAC	184384	Db	185405	CCCCCGGTACTGTACTCATGCCGCTGTGGCTTTTGGTGTCCCACTACTCGAGCCACC	185464	
Qy	2038	AACGACAGCTCATGGGCTTTGAGAGAGACTCTG-2072		Qy	2357	TGAAGCGGCACATGACAGACACACAGGGGGAGAGCGCTTCGCTGTGGCGCTGCCAT	2416	
Db	184385	AACGACAGCTCATGGGCTTTGAGAGAGACTCTG-2072		Db	185465	TGAAGCGGCACATGACAGACACACAGGGGGAGAGCGCTTCGCTGTGGCGCTGCCAT	185524	
Qy	2073	-----2072		Qy	2417	AGCCTCAGCCAGTTCGTCACCTGACCGGCACATACCCGACCCCATACTGCGGAGAGAC	2476	
Db	184444	AACGACAGCTCATGGGCTTTGAGAGAGACTCTGAGGTGTGTTAAGGTCTGGTTTA	184444	Db	185525	AGCCTCAGCCAGTTCGTCACCTGACCGGCACATACCCGACCCCATACTGCGGAGAGAC	185584	
Qy	2073	-----2072		Qy	2477	CTTACCGTTTCCCACTCCCTTTTTCCTGAGAGAGCTGGGCAACCTGAGGGCGGCATC	2536	
Db	184445	GGATTAGTGCAATTGGTAACTCAAGCCCTTTAGAAAGCGGTGCAGCGACCCGAAACAC	184504	Db	185585	CTTACCGTTTCCCACTCCCTTTTTCCTGAGAGAGCTGGGCAACCTGAGGGCGGCATC	185644	
Qy	2073	-----2072		Qy	2537	AGCGCACCCACACAGGGCTCCCACTCCCTGCGCCCACTGCTGGCTTTTCATGCTGTG	2596	
Db	184505	ACTTTTCAACCTAGAACTCCCTAAAGTCCCTCAGCAGCATCTTTTGTGAGGTGAA	184564	Db	185645	AGCGCACCCACACAGGGCTCCCACTCCCTGCGCCCACTGCTGGCTTTTCATGCTGTG	185704	
Qy	2073	-----2072		Qy	2597	CTCCAGCACCAACCCGCGCTCCCACTCCCAAGAGAGAGAGGGAACAATGCCCCGACAT	2656	
Db	184565	GTTCGGTTTCGGCTTAAGGTGTGTAATACAGGGGTGGGTAGCCCACTTCTATGCG	184624	Db	185705	CTCCAGCACCAACCCGCGCTCCCACTCCCAAGAGAGAGAGGGAACAATGCCCCGACAT	185764	
Qy	2073	-----2072		Qy	2657	CAGAAA-----2662		
Db	184625	CACCTCAGAGGTTTCATGCAAGGGAATTATGTAATAAGTACTCTTAAACTCCTTTT	184684	Qy	2663	-----2663		
Qy	2073	-----2072		Db	185765	CAGAAGGTAAAAAGAACACAGGCATCTTGGGACATGGGTGGTGTGATAGGATGCTT	185824	
Db	184685	GACCACTAGCAGACATCCTTTTGTACTGAGTTTCAGTTTCTATCCTTAAGTCTCTCA	184744	Qy	2663	-----2663		
Qy	2073	-----2072		Db	185825	GGATTATAGCCCATATCTCATTTTCCACAGATGCGCTGATCTCTGCCAGACTTGAATC	185884	
Db	184745	TTTATTTTGGGTAATTTTAAAGTCAGAGAAAAACCATATAAGAACTGCGCCTTTC	184804	Qy	2690	TTCTATGTCCACAGGTGTGTCAGTTTCTGTCAGACTGTGGGAGCTGCGGGGTGAAG	2749	
Qy	2073	-----2072		Db	185885	TTCTATGTCCACAGGTGTGTCAGTTTCTGTCAGACTGTGGGAGCTGCGGGGTGAAG	185944	
Db	184805	TGTGTAAATTGCACTACATGTTTCCACTCTCTTAATTTGCTTAAATGCTGCTTGAAGGCTT	184864	Qy	2750	GGGAGACTTGTGTGAATCTGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTGAACCT	2809	
Qy	2073	-----2072		Db	185945	GGGAGACTTGTGTGAATCTGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTGAACCT	186004	
Db	184865	TCATTTTCAATACCGCTTATATAAAAAAGGAGCTTACTATAACAGAGCTTACAAAAAG	184924	Qy	2810	CCCCGGGCTGTGACACAGAACTGAGAGAGGTGAGGAGAGCTGGGAGCTGCCATGT	2869	
Qy	2073	-----2072		Db	186005	CCCCGGGCTGTGACACAGAACTGAGAGAGGTGAGGAGAGCTGGGAGCTGCCATGT	186064	
Db	184925	ATCTTGGGATCTTATATTTTCTTGTCTCAAGAAATTAACGTTGCTTCAGCTCAGCT	184984	Qy	2870	GTGGGCGCTGCATCGAGGAGAGCTGGAGGGGTGCACTGGGGGACCCCGAGGGCCCTG	2929	
Qy	2073	-----2072		Db	186065	GTGGGCGCTGCATCGAGGAGAGCTGGAGGGGTGCACTGGGGGACCCCGAGGGCCCTG	186124	
Db	184985	GCACCTTACTTCAAGCTCATGATATTTTCCATTTGTTAAAAATGACATGCAACCTGGT	185044	Qy	2930	GTGACAAAGGCTTGTGCTGTAGTTTATGCCCCCTTTGGCACTACTACCCCAACACCTGG	2989	
Qy	2073	-----2072		Db	186125	GTGACAAAGGCTTGTGCTGTAGTTTATGCCCCCTTTGGCACTACTACCCCAACACCTGG	186184	
Db	185045	GTTAGTAAATTAGGCTCAGAGTAGGACCAAAATCAAGGATGAGGAGGTAAGTGGGCTA	185104	Qy	2990	CTCGGCATCAAGACTCAGAGTGTGAGAAACCTTTCCGCTGTGCCCGCTGTCCATACG	3049	
Qy	2073	-----2072		Db	186185	CTCGGCATCAAGACTCAGAGTGTGAGAAACCTTTCCGCTGTGCCCGCTGTCCATACG	186244	
Db	185105	CATGTGAGAGTGAAGTTTGGAACTCAGTCTTGGCCATGTTGACCGGTTGTTTTTTT	185164	Qy	3050	CCTCTGCTCATCTGGATAACCTGAACCGCACAGCGGTCCACACAGGAGAAAGCCCT	3109	
Qy	2073	-----AGGAGACTCTAGGGGGCAGACTGGACTTCCCTATGGGCTGAG	2117	Db	186245	CCTCTGCTCATCTGGATAAACCTGAACCGCACAGCGGTCCACACAGGAGAAAGCCCT	186304	
Db	185165	CCCTTACGTTTACACAGGAGACTCTAGGGGGCAGACCTGGACTTCCCTATGGGCTGAG	185224	Qy	3110	ACAGTGGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGGCTCATGGTC	3169	
Qy	2118	TGACGAGAGTCTGGGGCGCGCAGCTAAGTGGGAGAGTGAAGTTTGAAGAACAGC	2177	Db	186305	ACAGTGGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGGCTCATGGTC	186364	

QY	3170	GCATCCACTCTGGTGACAAACCTTTTCGGTGTAGCTTTTGCAACTACAGCTGCACACAGA	3229
Db	186365	GCATCCACTCTGGTGACAAACCTTTTCGGTGTAGCTTTTGCAACTACAGCTGCACACAGA	186424
QY	3230	GTATGAACCTCAAACCTGATATCTGCGGACACACCGGCGAGAGCCCTTCGGCTGTGCA	3289
Db	186425	GTATGAACCTCAAACCTGATATCTGCGGACACACCGGCGAGAGCCCTTCGGCTGTGCA	186484
QY	3290	CCTGGCCCTATACACAGGCCACTGGGACAACTACAAGGCTCATCAGAAGGTGCATGGCC	3349
Db	186485	CCTGGCCCTATACACAGGCCACTGGGACAACTACAAGGCTCATCAGAAGGTGCATGGCC	186544
QY	3350	ATGGTGGAGCAGAGGGCTGTGCTCTCTGCGCCCTGAGGGCTGGGCGCCACCTCATAGCC	3409
Db	186545	ATGGTGGAGCAGAGGGCTGTGCTCTCTGCGCCCTGAGGGCTGGGCGCCACCTCATAGCC	186604
QY	3410	CACCCCTCTGTTTGTAGCACTCGGGGTCACAGCCCTGGGTGCTACTGTGTAGCAGGGCTC	3469
Db	186605	CACCCCTCTGTTTGTAGCACTCGGGGTCACAGCCCTGGGTGCTACTGTGTAGCAGGGCTC	186664
QY	3470	TTCATTTCAGACTCACCTTGAACCTAGTGTCTCTGCGCCCTGAGGGCTGTAGGAATTAGCC	3529
Db	186665	TTCATTTCAGACTCACCTTGAACCTAGTGTCTCTGCGCCCTGAGGGCTGTAGGAATTAGCC	186724
QY	3530	CTATGCTCTGCTGATTTATACAACTGAACCTAGAAACCACTTCCCTTTCTCCCGCT	3589
Db	186725	CTATGCTCTGCTGATTTATACAACTGAACCTAGAAACCACTTCCCTTTCTCCCGCT	186784
QY	3590	GGTCAGGGGCTCCACACAGACTAAGCTAGGCACTATATGACAGAGCTGAATCCCATGT	3649
Db	186785	GGTCAGGGGCTCCACACAGACTAAGCTAGGCACTATATGACAGAGCTGAATCCCATGT	186844
QY	3650	CAGGGGCCATATAGACAGGGGACTGTCTTGTAGCTACGTACAGATGAGCTAAGTAT	3709
Db	186845	CAGGGGCCATATAGACAGGGGACTGTCTTGTAGCTACGTACAGATGAGCTAAGTAT	186904
QY	3710	TAGGGCTCTGATTCACCGCACTGTCTCCAGAGGCTATGATGAATCTGGAGCTG	3769
Db	186905	TAGGGCTCTGATTCACCGCACTGTCTCCAGAGGCTATGATGAATCTGGAGCTG	186964
QY	3770	CCCAGCCTTTTACTGTTTAACTTTTACGTGCTTTTATTAAGAAACATACAG	3829
Db	186965	CCCAGCCTTTTACTGTTTAACTTTTACGTGCTTTTATTAAGAAACATACAG	187024
RESULT 3	BD234722	2651 bp DNA linear	PAT 17-JUL-2003
LOCUS	BD234722	2651 bp DNA linear	PAT 17-JUL-2003
DEFINITION	BD234722	2651 bp DNA linear	PAT 17-JUL-2003
ACCESSION	BD234722	2651 bp DNA linear	PAT 17-JUL-2003
VERSION	BD234722.1	GI:33044492	
KEYWORDS	JP 2002516664-A/16		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Chen, Y. E., Horiuchi, M., Dzau, V. J., and Tamura, K.		
AUTHORS	Chen, Y. E., Horiuchi, M., Dzau, V. J., and Tamura, K.		
TITLE	Chen, Y. E., Horiuchi, M., Dzau, V. J., and Tamura, K.		
JOURNAL	Chen, Y. E., Horiuchi, M., Dzau, V. J., and Tamura, K.		
COMMENT	Chen, Y. E., Horiuchi, M., Dzau, V. J., and Tamura, K.		

QY	1198	ATGAGAGGGGAGAGGTACCACTTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	1257
Db	1	ATGAGAGGGGAGAGGTACCACTTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	60
QY	1258	ATTGGTTGGCTAGCAGCGCATCACAATCACTTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	1317
Db	61	ATTGGTTGGCTAGCAGCGCATCACAATCACTTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	120
QY	1318	TGTACAGCCCATGAACCTACGATCCCGTGTGTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	1377
Db	121	TGTACAGCCCATGAACCTACGATCCCGTGTGTCTCTGGTGTGGTGTGGTGTACCGCTCCAGCTTCTG	180
QY	1378	ACGAGGTACAGCTGTGCTTTCGAGAGGGGTAGGGGCTCCATGAATGAGAGCGGGC	1437
Db	181	ACGAGGTACAGCTGTGCTTTCGAGAGGGGTAGGGGCTCCATGAATGAGAGCGGGC	240
QY	1438	GGCGCGGGAGCGACCTGAGCTGGATTCGGGGGCGGGG-CAGGGGCTGCCAGGGGCGG	1496
Db	241	GGCGCGGGAGCGACCTGAGCTGGATTCGGGGGCGGGG-CAGGGGCTGCCAGGGGCGG	300
QY	1497	CACCGTGTATGGGGCGGTTCTGTGATCTTAAGAGCAAGGACCGACGGCGCGAACT	1556
Db	301	CACCGTGTATGGGGCGGTTCTGTGATCTTAAGAGCAAGGACCGACGGCGCGAACT	360
QY	1557	GGAGGGGCGGGGCGGGAGGCGCTCGGGCGGGGCGGGG-CAGGGGCTGCCAGGGGCGG	1616
Db	361	GGAGGGGCGGGGCGGGAGGCGCTCGGGCGGGGCGGGG-CAGGGGCTGCCAGGGGCGG	420
QY	1617	TGCCCCGGAAGCAAGCAAGCCATCCAGAGCCCGTGAATTCGAGGGGCTCAAGGTGAGG	1676
Db	421	TGCCCCGGAAGCAAGCAAGCCATCCAGAGCCCGTGAATTCGAGGGGCTCAAGGTGAGG	480
QY	1677	GGTCAGGGGCTTGAAGCGGGAGGCAAGGGGTGGGTCAGTAGTGGGCTCAGGTGAG	1736
Db	481	GGTCAGGGGCTTGAAGCGGGAGGCAAGGGGTGGGTCAGTAGTGGGCTCAGGTGAG	540
QY	1737	GGTCAGGGGCTTGAAGCGGGAGGCAAGGGGTGGGTCAGTAGTGGGCTCAGGTGAG	1796
Db	541	GGTCAGGGGCTTGAAGCGGGAGGCAAGGGGTGGGTCAGTAGTGGGCTCAGGTGAG	600
QY	1797	AGTATTAAGAAACCTGAGGGTTCAGTAGTACGGAAGTGCAGTTCACACAGCTGGCTC	1856
Db	601	AGTATTAAGAAACCTGAGGGTTCAGTAGTACGGAAGTGCAGTTCACACAGCTGGCTC	660
QY	1857	CTTGGTTCGATTATGGGTACTGCTTGGAGGCGAGATTCCACAGCACCCCTCCCTCTTT	1916
Db	661	CTTGGTTCGATTATGGGTACTGCTTGGAGGGAGATTCCACAGCACCCCTCCCTCTTT	720
QY	1917	AGTGGTACTGAAGATTCCTTCGAGGAAGGTCTCGGGGCTTGGTGTGGAGAGCGATT	1976
Db	721	AGTGGTACTGAAGATTCCTTCGAGGAAGGTCTCGGGGCTTGGTGTGGAGAGCGATT	780
QY	1977	GCTACTAGCCCAAGTCTGGAGTTTGAAGAGGAAGAGGAGGATGAAGGTGACGGCCA	2036
Db	781	GCTACTAGCCCAAGTCTGGAGTTTGAAGAGGAAGAGGAGGATGAAGGTGACGGCCA	840

AS1K37/02,
PC A61K37/48, C12N5/00
CC CNRE binding factors and uses thereof
FH Key Location/Qualifiers
FT source 1.2651
PT /organism='Mus musculus (mouse)'.
FEATURES
source
Location/Qualifiers
1.2651
/organism='Mus musculus'
/mol_type='genomic DNA'
/db_xref='taxon:10090'

Query Match 68.3%; Score 2628.4; DB 6; Length 2651;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2643; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

2037 CAACGACGAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC 2096
 Db
 841 CAACGACGAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC 900
 Qy
 2097 TGGACTTCCCTATGGGCTGAGTGACGAGAGTCTGGGGGGCCGCGCACATAAGTGCGGA 2156
 Db
 901 TGGACTTCCCTATGGGCTGAGTGACGAGAGTCTGGGGGGCCGCGCACATAAGTGCGGA 960
 Qy
 2157 GAGTGAAGTTGAGGAACACAGCCAGAGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGAGGCC 2216
 Db
 961 GAGTGAAGTTGAGGAACACAGCCAGAGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGAGGCC 1020
 Qy
 2217 AGCCTGTCTGAGTGTGTGGGGGCCGAGCAAGGTGAGGGGCCGTGTGTGTGGGGCAGAGGGCG 2276
 Db
 1021 AGCCTGTCTGAGTGTGTGGGGGCCGAGCAAGGTGAGGGGCCGTGTGTGTGGGGCAGAGGGCG 1080
 Qy
 2277 GGGTGGGGGGCCCGCTGCCCTCAGCGTTACTGTACTCATGCGGCTGTGGCTTTTGGT 2336
 Db
 1081 GGGTGGGGGGCCCGCTGCCCTCAGCGTTACTGTACTCATGCGGCTGTGGCTTTTGGT 1140
 Qy
 2337 GTCCCACTACTCGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGGAGAACCGTT 2396
 Db
 1141 GTCCCACTACTCGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGGAGAACCGTT 1200
 Qy
 2397 CGGCTGTGCGCCCTGCCATAGGNTCAGCCAGTTCGTCAACCTGACGCGACATACCGG 2456
 Db
 1201 CGGCTGTGCGCCCTGCCATAGGNTCAGCCAGTTCGTCAACCTGACGCGACATACCGG 1260
 Qy
 2457 CACCCATATCTGGCGAGAACCCCTACGTTGTGCCCACTGCCCTTTGGCTGACGAGCGCT 2516
 Db
 1261 CACCCATATCTGGCGAGAACCCCTACGTTGTGCCCACTGCCCTTTGGCTGACGAGCGCT 1320
 Qy
 2517 GGGGAACCTGAGCGGCTCAGCGGACCCACACAGAGGCTCCACTCTCCCTGCCCAAC 2576
 Db
 1321 GGGGAACCTGAGCGGCTCAGCGGACCCACACAGAGGCTCCACTCTCCCTGCCCAAC 1380
 Qy
 2577 CTGTGGCTTTTCAGTGTGTGCTCCAGCAACCCGGCTCCAGTCCCAACAGAGCAGGA 2636
 Db
 1381 CTGTGGCTTTTCAGTGTGTGCTCCAGCAACCCGGCTCCAGTCCCAACAGAGCAGGA 1440
 Qy
 2637 GGGGCAATATGCCCGGACGATCAGAAAATGCGTGATCTGCGCAGACTTGAAGTCTTCATGT 2696
 Db
 1441 GGGGCAATATGCCCGGACGATCAGAAAATGCGTGATCTGCGCAGACTTGAAGTCTTCATGT 1500
 Qy
 2697 GCACACAGTGTGTGCGAGTTCTCTGACAGCTGTGGGAGCTGCGGGGTGAAGGGGAGAG 2756
 Db
 1501 GCACACAGTGTGTGCGAGTTCTCTGACAGCTGTGGGAGCTGCGGGGTGAAGGGGAGAG 1560
 Qy
 2757 CTGTGTGGAATGGATCCGAACCACTGCGCAGAGCTACTGTTCCCTTGGACCTGCCGGG 2816
 Db
 1561 CTGTGTGGAATGGATCCGAACCACTGCGCAGAGCTACTGTTCCCTTGGACCTGCCGGG 1620
 Qy
 2817 CTGTGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2876
 Db
 1621 CTGTGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
 Qy
 2877 CTGCGATGCGAGGAGAGGCTGGAGGGGTTCGCCACTGCGGGACCCAGGGCCCTGGTGACAA 2936
 Db
 1681 CTGCGATGCGAGGAGAGGCTGGAGGGGTTCGCCACTGCGGGACCCAGGGCCCTGGTGACAA 1740
 Qy
 2937 AGGCTTTGGCTGTAGTTATGCGCTTTGCGACTCACTACCCCAACCACTGGCTGGCGCA 2996
 Db
 1741 AGGCTTTGGCTGTAGTTATGCGCTTTGCGACTCACTACCCCAACCACTGGCTGGCGCA 1800
 Qy
 2997 CATGAAGACTCAAGTGGTGAGAAAACCTTCCGCTGTGCCCTGTCCATAGCCCTCTGC 3056
 Db
 1801 CATGAAGACTCAAGTGGTGAGAAAACCTTCCGCTGTGCCCTGTCCATAGCCCTCTGC 1860
 Qy
 3057 TCATCTGGATTAACCTGAAGCGGACAGCGGCTCCACAGAGAGAAAAGCCCTACAAGTG 3116
 Db
 1861 TCATCTGGATTAACCTGAAGCGGACAGCGGCTCCACAGAGAGAAAAGCCCTACAAGTG 1920
 Qy
 3117 CCCCCTCTGCTGCTGTGTGGCAACCTGGGCCAACCTCAAGCGCTCATGGTCCGATCCA 3176

1921 CCCCCTCTGCTGCTGATGCTGTGTGGCAACCTGGCAACCTCAAGCGCTCATGGTCCGATCCA 1980
 Qy
 3177 CTCTGCTGACAAACCTTTTTCGGTGTAGCCTTTTTCGAACCTACAGCTGCAACCCAGAGTATGAA 3236
 Db
 1981 CTCTGCTGACAAACCTTTTTCGGTGTAGCCTTTTTCGAACCTACAGCTGCAACCCAGAGTATGAA 2040
 Qy
 3237 CTTCAAAACCTCATATCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCGCACCTGCGC 3296
 Db
 2041 CTTCAAAACCTCATATCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCGCACCTGCGC 2100
 Qy
 3297 CTATACACAGGGCCACTGGGACAACTACAAGCGTCTATAGAAGTGCATGCGCATGGTGG 3356
 Db
 2101 CTATACACAGGGCCACTGGGACAACTACAAGCGTCTATAGAAGTGCATGCGCATGGTGG 2160
 Qy
 3357 AGCAGAGGGCCCTGCTCTCTGCCCTCAGGGCTGGGCGCCACCTCATAGCCACCTC 3416
 Db
 2161 AGCAGAGGGCCCTGCTCTCTGCCCTCAGGGCTGGGCGCCACCTCATAGCCACCTC 2220
 Qy
 3417 TGTGTTGAGCACTCGGGTCCAGCAGCCCTGGGTGTACTGTGAGCAGGGCTCTTCAATTC 3476
 Db
 2221 TGTGTTGAGCACTCGGGTCCAGCAGCCCTGGGTGTACTGTGAGCAGGGCTCTTCAATTC 2280
 Qy
 3477 AGACTCACTTTGAACCTAGGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 3536
 Db
 2281 AGACTCACTTTGAACCTAGGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 2340
 Qy
 3537 TCTGCACTTTTATACAAATGAACCTAGAACCACTTTTCCCTTCTCCCGCTGGTCAGG 3596
 Db
 2341 TCTGCACTTTTATACAAATGAACCTAGAACCACTTTTCCCTTCTCCCGCTGGTCAGG 2400
 Qy
 3597 GGCTCCACACAGACTAACCTAGGCACTATATGACAGCCTGAATCCCATGGTCAGGGG 3656
 Db
 2401 GGCTCCACACAGACTAACCTAGGCACTATATGACAGCCTGAATCCCATGGTCAGGGG 2460
 Qy
 3657 CCATATAGACAGGGGACTTGTCTTAGCTCACGTACAGATGAGCTAAGTGAATAGGGCC 3716
 Db
 2461 CCATATAGACAGGGGACTTGTCTTAGCTCACGTACAGATGAGCTAAGTGAATAGGGCC 2520
 Qy
 3717 TTGGATTACCGGCACTGCTCCAGAGGCTATGATGAAGTGGTGGAGGCTGCCAGCC 3776
 Db
 2521 TTGGATTACCGGCACTGCTCCAGAGGCTATGATGAAGTGGTGGAGGCTGCCAGCC 2580
 Qy
 3777 TTTTACTGTTTTAACTATTATTCTAGTGTCTTATATAAGGAAACACTAACAGAAAAAAA 3836
 Db
 2581 TTTTACTGTTTTAACTATTATTCTAGTGTCTTATATAAGGAAACACTAACAGAAAAAAA 2640
 Qy
 3837 AAAAAAAAAA 3847
 Db
 2641 AAAAAAAAAA 2651

RESULT 4

BD234713
 LOCUS 2289 bp DNA linear PAT 17-JUL-2003
 DEFINITION CNRE binding factors and uses thereof.
 ACCESSION BD234713
 VERSION BD234713.1 GI:33044483
 KEYWORDS JP 2002516664-A/7.
 SOURCE Mus musculus (house mouse).
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2289)
 AUTHORS Chen, Y. E., Horiuchi, M., Dzu, V. J. and Tamura, K.
 TITLE CNRE binding factors and uses thereof
 JOURNAL Patent: JP 2002516664-A 7 11-JUN-2002;
 THE BRIGHAM AND WOMEN'S HOSPITAL INC
 COMMENT OS Mus musculus (mouse)
 PN JP 2002516664-A/7
 PD 11-JUN-2002
 PF 23-APR-1999 JP 2000545541
 PR 24-APR-1998 US 60/082997

PI	YUQING E CHEN, MASATSUGU HORIUCHI, VICTOR J DZAU, KOICHI TAMURA
PC	C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K45/06,
PC	A61K48/00,
PC	A61P9/00, A61P9/04, A61P9/12, A61P43/00, C07K4/47, C12N1/15, C12N1/
PC	19, C12N1/21,
PC	C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, C12N15/00, PC
PC	A61K37/02,
PC	A61K37/48, C12N5/00
CC	CNRE binding factors and uses thereof
PH	Key Location/Qualifiers
FT	source 1. .2289
FT	Location/Qualifiers
FEATURES	
source	1. .2289
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
ORIGIN	
	Query Match 59.5%; Score 2288; DB 6; Length 2289;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1198 ATGAGGGGGCAGGGTACCACTTCCTCGGTGGCTGGTTGGGTACGCTCCAGCTTCG 1257
Dd	1 ATGAGGGGGCAGGGTACCACTTCCTCGGTGGCTGGTTGGGTACGCTCCAGCTTCG 60
Qy	1258 ATTGGTTGGCTAGCAGCGCATCAATCACTTCCTGCTCATGCTGGACTGGCTCCGGT 1317
Dd	61 ATTGGTTGGCTAGCAGCGCATCAATCACTTCCTGCTCATGCTGGACTGGCTCCGGT 120
Qy	1318 TGTACAGCCCATGAATPACGCATCCCGTGTGCTCTCGGCTGGCGGAAGCGAAGCGGGT 1377
Dd	121 TGTACAGCCCATGAATPACGCATCCCGTGTGCTCTCGGCTGGCGGAAGCGAAGCGGGT 180
Qy	1378 ACGAGGTACGAGTGTCTTTCGAGGGGGGTAGGGGCTTCATGATGAGACGCGCGGC 1437
Dd	181 ACGAGGTACGAGTGTCTTTCGAGGGGGGTAGGGGCTTCATGATGAGACGCGCGGC 240
Qy	1438 GCGCGCGGAGCGACTGAGTGGATTCCGGGGCCGGGCGAGGGCTGCCACGGGCCCGC 1497
Dd	241 GCGCGCGGAGCGACTGAGTGGATTCCGGGGCCGGGCGAGGGCTGCCACGGGCCCGC 300
Qy	1498 ACCGTGTATGGGGCGGTTCTGTGATCCTAAGAGCAAGGACCGACGCGAGGGCCGAAC 1557
Dd	301 ACCGTGTATGGGGCGGTTCTGTGATCCTAAGAGCAAGGACCGACGCGAGGGCCGAAC 360
Qy	1558 GGAGGGCGGGCCGGGAGCGCTCGGCGCGAGCGCGCTGGGCTGGAGCCGCTCAGCAT 1617
Dd	361 GGAGGGCGGGCCGGGAGCGCTCGGCGCGAGCGCGCTGGGCTGGAGCCGCTCAGCAT 420
Qy	1618 GCCCCGAAGGAAGCAAAACCATCCACAGCCGTTGAAATGCGAGGGGCTCAAGGTCAGGG 1677
Dd	421 GCCCCGAAGGAAGCAAAACCATCCACAGCCGTTGAAATGCGAGGGGCTCAAGGTCAGGG 480
Qy	1678 GTACGGGCGCTTGAGCCGGGAGGAACAGGGGTGGGTCAGTAGNGTGGGCTCAGGTCAGG 1737
Dd	481 GTACGGGCGCTTGAGCCGGGAGGAACAGGGGTGGGTCAGTAGAGTGGGCTCAGGTCAGG 540
Qy	1738 GTGAGGGGGACTCCTCAGGGTTAGGGCGCGATGATCTGGGATCTTCGCTCCCTTACCAGA 1797
Dd	541 GTGAGGGGGACTCCTCAGGGTTAGGGCGCGATGATCTGGGATCTTCGCTCCCTTACCAGA 600
Qy	1798 GTATTAAAGAACTTGAGGGTCAATCGAGTACGGGAATGCAAGTTCACACAGCTGGCTCC 1857
Dd	601 GTATTAAAGAACTTGAGGGTCAATCGAGTACGGGAATGCAAGTTCACACAGCTGGCTCC 660
Qy	1858 TTGGTTCGGATTATGGGTACTGCTTGGAGGAGGATTCACAAAGCCCTCCCTCTTTA 1917
Dd	651 TTGGTTCGGATTATGGGTACTGCTTGGAGGAGGATTCACAAAGCCCTCCCTCTTTA 720
Qy	1918 GTGGATCTGAAGATTCCTTCGACGAAGGTCCTGGGGCCCTGGTGTGGAGAGCGGATTCG 1977

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 270488: contig of 270488 bp in length
 * 270489: gap of unknown length
 * 270589: contig of 1405 bp in length
 * 271994: gap of unknown length
 * 272094: contig of 1967 bp in length.

FEATURES

source
 1. .274060
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-9D7"
 /cdate="1100"
 /notes="wgs_contig"
 214108. 216519
 /notes="wgs_contig"
 268481. 270488
 /notes="wgs_contig"



ORIGIN

Query Match 54.4%; Score 2092; DB 2; Length 274060;
 Best Local Similarity 73.5%; Pred. No. 0;
 Matches 3373; Conservative 0; Mismatches 315; Indels 902; Gaps 22;

Qy 124 CAGNAAATTTTTCAGATGTGACGAGAGTGGAGTAGTACCT--GAAACAT 181
 Db 265200 CCACACATCTCTCTGTATGACACGAGGCTGGGAAATAGTGTAAATAGGACACAC 265141

Qy 182 TAGTGACTTTGATAAAGGNGGTTTTCCTTCATTTCCAGATGTAG-CCTGGGCTTC 240
 Db 265140 TAGAAGACCTGATTAAGAGATGAATCTTCCTGTTTCCCTGATACAGTCTTAGACTC 265081

Qy 241 AGAAATNAC----NGTTTTTTTTTGATTAACCTTGCCATCTTTGTGCTCTT 296
 Db 265080 AGAAATTCAAAAATTTCTGTTTGTGATAATTTGTGGCCATCTCTCTCTGCCAC 265021

Qy 297 CTGTATTACAGGACAACTTTCTGTGTTACTTT-----GGTCAAAC 342
 Db 265020 CATGTATCAGACAGGACAACTTTCTATGTTACTCTGTGTTCCCGGAAGTGTCAA 264961

Qy 343 CCATGGCTAACTGCTTAACTTTTGTAGCTGTGATCAGTCTGCGGCCAGACA 402
 Db 264960 CTTGTGGCTAACTTCACTCAACCTTTTGTAGCTGTGATCAGTCTGCGGCCAGACA 264901

Qy 403 CTTCTGGGATGTACAGGCTGTGACAACTGATGATCATCTTTTCAAGCCCC 462
 Db 264900 CTTCTGGGATGTGACAGGCTGTGACAACTGATGATCATCTTTTCAAGCCCC 264841

Qy 463 GAAACACAGTAGAGCTTCAGGCGAGAGTGGCAAGAGGAACTGGAGAGGCACTGTCCA 522
 Db 264840 GAAACACAGTAGAGCTTCAGGCGAGAGTGGCAAGAGGAACTGGAGAGGCACTGTCCA 264781

Qy 523 CGAGGGGCTGAGACACCGGCAAGTACAAAGAGAGGCAAGAGGATAGTGT 582
 Db 264780 CGAGGGGCTGAGAGATGGCAAGTACAAAGAGAGGCAAGAGGATAGTGT 264721

Qy 583 CAACCGGACCTGCGCACTGTGAGTGTCTTCTGAGCCCTTGGACCGGAGCTGAGTTTG 642
 Db 264720 CAACCGGACCTGCGCACTGTGAGTGTCTTCTGAGCCCTTGGACCGGAGCTGAGTTTG 264663

Qy 643 TCCTTGTCTTTAGCTTACAGTGGGTATGAGTGTGAGGGGCTGGGTGCTTTCCT 702
 Db 264662 TCCTTGTCTTTAGCTTACAGTGGGTATGAGTGTGAGGGGCTGGGTGCTTTCCT 264603

Qy 703 CAGGCCATTACAAAGAGGCCCCCACCCTCCCGGAGGCTGGAGGCTCTGCTG 762
 Db 264602 CAGGCCATTACAAAGAGGCTCT---CCCTTGGACCGGAGGCTGGAGGCTCTGCTG 264546

Qy 763 TCCTCTTAAGCCCTCTTACT-CTCCTTGGGGCTCATGACTATGCGTCTGTGCTGTGCT 821

Db 264545 TCCTCTCAGCCCTCCTTACTCTCTTGGGCTCATGACTATCGGTTCTGTGCTGTGCT 264486
 Qy 822 CTGTGTGTTGGAGGAGGAGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 881
 Db 264485 CTGTGTGTTGGAGGAGGAGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 264426

Qy 882 AGGACATCTTTTATTTGCGGCTCTGTGACCCCTAGCCGCTTGCACCCGCTCTCTGTT 941
 Db 264425 AGGACATCTTTTATTTGCGGCTCTGTGACCCCTAGCCGCTTGCACCCGCTCTCTGTT 264366

Qy 942 GTACACTTCAAGCAACACATTTTTCAGACTAAAGGCGCAAAACAAAGCTAATCTGTCTCAT 1001
 Db 264365 GTACACTTCAAGCAACACATTTTTCAGACTAAAGGCGCAAAACAAAGCTAATCTGTCTCAT 264306

Qy 1002 AGTGTCTATCTTTTACTCTCTACCTGCGCCCATATGTGTGTGCGCGCTGTGTGTGCTG 1061
 Db 264305 AGTGTCTATCTTTTACTCTCTACCTGCGCCCATATGTGTGTGCGCGCTGTGTGTGCTG 264246

Qy 1062 CTTGTGTGTCTCCGAGCCTCTGGAGAGGGGCGAGTGAATGTGGAGGAGGAGCTGGAGT 1121
 Db 264245 --TGTGTGAGTCTCTGAGCCTCTGGAGGGGAG-----CAGTGTGGATCAGGAGCTGGAAGT 264193

Qy 1122 ----CTTAACCTCTGAGTAAAGCAGCCGCTGGAGGCCATCCT-GCGCTGTTCGCTTCAG 1176
 Db 264192 AGTGTAACTCTTAAAGCAGCAGCCGAGAGGGCCATCTGTGCGTGTTCGCTTCAG 264133

Qy 1177 GCACGACTTCTTACTGCTCT-ATATGGAGGGGCGAGGTACACCTTCTGTTGGCTGCT 1235
 Db 264132 GCATGACTTCTTACTGCTCT-ATATGGAGGGGCGAGGTACACCTTCTGTTGGCTGCT 264073

Qy 1236 TGGGTACGCTCAGCTTCTGATTTGTTGGCTAGCAGCGCATCAATCACTTCTCTGCT 1295
 Db 264072 CGGCTCATTTTATTCAGCTTCTTATGGCCGCTAGCAGCGCATCAATCACTTCTCTGCT 264013

Qy 1296 CATGCTGAGCTGCGCTCGGCTTGTACAGCCCATGAACTAGCATCCCGTGTCTCTGCT 1355
 Db 264012 CATGCTGAGCTGCGCTCGGCTTGTACATATTCATTAAGTCTTCCGAGTCTCTGCT 263954

Qy 1356 GGTGCGGAAAGCGGAGCGGTACGAGGTACAGCTGTGTTCTTGGAGGGGGGTAGGGGG 1415
 Db 263953 GGTGCGGAAAGCGGAGCGGTACGAGGTACAGCTGTGTTCTTGGAGGGGGGTAGGGGG 263894

Qy 1416 CTCATGATTAAGTGAAGCGGCGCGCGGAGCGAGCTGAGCTGGATTCGCGGGGCC--GG 1474
 Db 263893 CTCATGATTAAGTGAAGCGGCGCGCGGAGCGAGCTGAGCTGGATTCGCGGGGCCAGG 263834

Qy 1475 GGCAGGGCTGCCCAAGGCGCGCACCC-GTGTATGGGGCGGTTCTGTGATCTTAAGAGCA 1533
 Db 263833 GGCAGGGCTGCCCAAGGCGCGCACCCGTTGTATGGGGCGGCTCTGTGACCTTAAGAGCA 263774

Qy 1534 AGGACCGAGCGAGCGCGCGGAGCGCGCGGAGCGCTCGGAGCGCTCGGCGCGGAGCG 1593
 Db 263773 AGGACCGAGCGGCTGGGCGGATCTGGAGGGGCGGCGCGGAGCGCTCGGCGCGGAGCG 263714

Qy 1594 CCGTGGGCTGGAGCGCGCTCACGATGCCCGGAGGAGCAAAAGCCATCCACAGCCGCTGAA 1653
 Db 263713 CCGTGGGCTGGAGCGCGCTCACGATGCCCGGAGGAGCAAAAGCCATCCACAGCCGCTGAA 263654

Qy 1654 ATGCCAGGGGTCAAGGCTCAGGGCTCAGGGCTTGGAGCGGGAGGAGCAAGGGTGGGG 1713
 Db 263653 ATGCCAGGGGTCAAGGCTCAGGGCTCAGGGCTTGGAGCGGGAGGAGCAAGGGTGGGG 263594

Qy 1714 TCAGTAGAGTGGGCTCAGGCTCAGGCTGAGGGGAGCTCCTCAGGGTTAGGGCGGATGAT 1773
 Db 263593 TCAGTAGAGTGGGCTCAGGCTCAGGCTGAGGGGAGCTCCTCAGGGTTAGGGCGGATGAT 263534

Qy 1774 CTGGGATCTTCTGCTCTTACCGAGATTAAGAGAACCTGAGGCTCATCGAGTACGGAA 1833
 Db 263533 CTGGGATCTTCTGCTCTTACCGAGATTAAGAGAACCTGAGGCTCATCGAGTACGGAA 263474

Qy 1834 GTGAGTTTCAACAGCTGGCTCTCTGTTGTTATGAGTTATGGGTACTGCTTGGAGGAGAGAT 1893

Db 263473 GTGCAGTTTCAACAGCTGGCTCCCTGTTTTCGAATTATATGGAAACGGCTTTGAGGGAGAT 263414
 Qy 1894 TCCACAGACCCCTCCCTCTTTAGTGGATCTGATGATTTCTTTCAGCAAGATCTCTGG 1953
 Db 263413 TCCACAGACCCCTCCCTCTTTAGTGGATCTGATGATTTCTTTCAGCAAGATCTCTGG 263354
 Qy 1954 GCCCTGTTTGGAGACGATTTGCTACTAGGCCAAGATCTGAGTTTGAAGAGGAAG 2013
 Db 263353 GCTCTGTTTGGAGATGATTTGCTACTAGGCCAAGATCTGAGTTTGAAGAGGAAG 263294
 Qy 2014 GAAGAGATGAGTGGAGGCCCAACAGCAGCTCATGAGCTTTGAGAGAGACTCTGA- 2072
 Db 263293 GAAGAGAGGAAGTGAAGGCCCAACAGCAGCTCATGAGCTTTGAGAGAGACTCTGA 263234
 Qy 2073 ----- 2072
 Db 263233 GGTGTTTAAAGGTCGTGGTTAAGGATAAGAGGCTTTTGGTAACTCAAGCCCTTTAGAA 263174
 Qy 2073 ----- 2072
 Db 263173 GCGGTTTGGAGACCAAAACACCACTTTTACAACTAGCAGCAGCATTTTGGTGA 263114
 Qy 2073 ----- 2072
 Db 263113 GTTGAAGTTCTGGGTCGGGCTTAAGGTGGTGTGAATACAGAGGTTGGTAGCCACATTC 263054
 Qy 2073 ----- 2072
 Db 263053 TTATGTACCCCTCAAGAGGTTTCATGGAAGGAATATGTAATAAGTACTCTTAAAACT 262994
 Qy 2073 ----- 2072
 Db 262993 CTCTGGCTAGTACGATAGATCTTTTGTGTTCTGAGGTTTCAGTTTCTGCCCTTAACT 262934
 Qy 2073 ----- 2072
 Db 262873 CTCTCTGTTTATTTCTTGGTATTTTAAAGTCAGAGAAACCATATAAAGAACTCTGC 262874
 Qy 2073 ----- 2072
 Db 262873 CCTTCTGTGTAATTTGGCACTACATGCTCCGACTTCTCCAAATTTGCTACTTCTGAGA 262814
 Qy 2073 ----- 2072
 Db 262813 TGGCTTTTCAATATACAGCTTTATATAAAGAGGCTTTACTAAACAGGGTTTACA 262754
 Qy 2073 ----- 2072
 Db 262753 AAAAGATCTTGGGTCCTTATATTTTACTTATCTCATATTTGCTTTCACATCAGCCT 262694
 Qy 2073 ----- 2072
 Db 262693 GTACTCCACTTCAAGCCTGTGATATTTCCATTTGTTGAAATGATTAAGTCAACCCCTAGT 262634
 Qy 2073 ----- 2072
 Db 262633 GTTAGTAATTAGGCTCAATAGATGAATCAAGGTGAGCAGAAATAGTGGGCTTTT 262574
 Qy 2073 ----- 2072
 Db 262573 TGTGAGAAATGAGATTTGGAATCTGATGAGCCATGTTGATTTGTTTCTTTCTTTT 262514
 Qy 2073 ----- AGAGACTCTCAGGGGCCAGACC 2096
 Db 262513 TTTTCTTTTTCCTCCCTCTCTTTTCTGTTTCAACAGAGACTCTCAGGGAGCCAGACC 262454
 Qy 2097 TGGACTTCCCTATGGGCTCAGTGAGCAGAGTCTGGGGCGGCCCGCACTAAAGTGGCGA 2156
 Db 262453 TGGACTTCCCTATGGGCTCAGTGAGCAGAGTCTGGGGCGGCCCGGCACTAAAGTGGCGA 262394
 Qy 2157 GAGTGAAGTTGAGGAACCAAGCCAGGGGTCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC 2216
 Db 262393 AAGTGAAGTTGAGGAGCCAGCGAGGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC 262334

Qy 2217 AGCTGTCTAGTGTGTGGGGGCGCAGCAGGTGAGGGGCGGTGTGTGGGGCAGAGGGCG 2276
 Db 262333 AGCTGTCTAGTGTGTGGGGGCGCAGCAGGTGAGGGGCGGTGTGTGGGGCAGAGGGCG 262274
 Qy 2277 GGTGGG-----GGGCCCCCGCTGCCCCCAGGTACTGTACTCATGCGGCTGTGC 2328
 Db 262273 GGTGGGNN 262214
 Qy 2329 GCTTTCTGTCTCCACTACTCTAGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGAG 2388
 Db 262213 GCTTTCTGTCTCCACTACTCTAGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGAG 262154
 Qy 2389 AGCGGTTCCGCTGGCGGTGCCATACGCTCAGCCAGTTCTGTCACCTCAGCGCA 2448
 Db 262153 AAGCGTTCCGCTGGCGGTGCCCTACGCTCAGCCAGCTCTGTCACCTCAGCGCA 262094
 Qy 2449 CATACCGCACCCTACTTGGCGAGAGCCCTACGTTTGTCCCCACTGCCCCCTTTGCCCTGC 2508
 Db 262093 CATACCGCACCCTACTTGGCGAGAGCCCTACGCTGTCCCCACTGCCCCCTTTGCCCTGC 262034
 Qy 2509 AGAGGCTGGGCACTCTAGCGGCGATCAGCGCACCACAGAGGCCCTCCCACTCTCTCC 2568
 Db 262033 AGAGGCTGGGCACTCTAGCGGCGATCAGCGCACCACAGAGGCCCTCCCACTCTCTCC 261974
 Qy 2569 TGCCTCAACTGTGGGTTTTCGATCTGTGTCTCCAGCAACCCCGGCTCCCACTCTCCACA 2628
 Db 261973 TGCCTCAACTGTGGTTTTCGATCTGTGTCTCCAGCAACCCCGGCTCCCACTCTCCACA 261914
 Qy 2629 GAGCAGGGGAGCAATGCCCCCGAGCATCA----- 2660
 Db 261913 GAGCAGGAGGAGCAATGCCCCCGGCGATCAGAGGTAAGAGCAGCAGAGATCTTGGG 261854
 Qy 2661 -----A 2661
 Db 261853 AACTGTAGTAGCAGTAGGATACTTGGATTAATGGCCCATATCTCATTTGTTCCACA 261794
 Qy 2662 AATGCCCTGATCTGCGCACTTGTGATCTTCTATGTCGCAAGAGTGTGCTGCTTCTG 2721
 Db 261793 GATGCTCTCATCTCCAGACCTGATCTTCCATGTGCCACAGGTGTGATCCAGTTTCTG 261734
 Qy 2722 CCAGACTGTGGGAGCTGCGGGTGAAGGGGAGAGCTTGTGTGGAAGTGGATCCGAACCA 2781
 Db 261733 CCAGATTGTGGCAGCTGCGGGTGAAGGGGAGGGTTTGTGTGGAAGTGGATCAGAACCA 261674
 Qy 2782 CTGCCAGACTACTGTTTCCCTTGGACCTCCGGGCTGTGGAAGGAACTGGAGAGGGT 2841
 Db 261673 CTGCCAGACTACTGTTTCCCTTGGACCTCCGGGCTGTGGAAGGAACTGGAGAGGGT 261614
 Qy 2842 GAGGCGAGAGGCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2901
 Db 261613 GAGGCGAGAGGTTGGGAACTGCCATGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 261554
 Qy 2902 GTTGGCACTGGGGGACCCAGGGGCTTGTGCAAAAGGCTTTTGTGCTGCTGCTGCTGCTG 2961
 Db 261553 GGTGGCAGTGGGGGCCCCCAGGGGCTTAGTGCAAAAGGCTTTTGTGCTGCTGCTGCTGCTG 261494
 Qy 2962 TTTGCCACTCACTACCCCAACCACTGCTCGGCACTAGAGCTCAGAGTGTGTGAGAA 3021
 Db 261493 TTTGCTACTCACTACCCCAACCACTGCTCGGCACTAGAGCTCAGAGTGTGTGAGAA 261434
 Qy 3022 CCTTTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3081
 Db 261433 CCTTTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 261374
 Qy 3082 CAGCGGTCCACAGAGAGAAAGCCCTACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3141
 Db 261373 CAGCGGTCCACAGAGAGAAAGCCCTACAGTGCCTCTCTGCTGCTGCTGCTGCTGCTGCTG 261314
 Qy 3142 AACTGCCCAACTCAAGCGTCTATGTCGATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 3201
 Db 261313 AACTGCCCAACTCAAGCGTCTATGTCGATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 261254

Db 841 CAAAGCAGCTCATGTGGCTTTGAGAGAGACTCTGAAGGAGACTCTCAGGGGCCAGACC 900
QY 2097 TGGACTTCCCTATGCGGTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAGTGCAGA 2156
Db 901 TGGACTTCCCTATGCGGTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAGTGCAGA 960
QY 2157 GAGTGAAGTTGAGAAACAGCAGGAGTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCC 2216
Db 961 GAGTGAAGTTGAGAAACAGCAGGAGTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCC 1020
QY 2217 AGCTGTGAGTGTGGGGGGCGCAGAGTGAAGGGGCGGTGTGGGGCGAGAGGGCG 2276
Db 1021 AGCTGTGAGTGTGGGGGGCGCAGAGTGAAGGGGCGGTGTGTGGGGCGAGAGGGCG 1080
QY 2277 GGGTGGGGGGCGCGCTGCCCTCCACCGTTACTGTACTCATGCGCGGTGTGGCGTTTGGT 2336
Db 1081 GGGTGGGGGGCGCGCTGCCCTCCACCGTTACTGTACTCATGCGCGGTGTGGCGTTTGGT 1140
QY 2337 GTCCCACTACTCGAGCCACTGAGCGGCACATGACAGACACAGCGGGGAGAGCGGTT 2396
Db 1141 GTCCCACTACTCGAGCCACTGAGCGGCACATGACAGACACAGCGGGGAGAGCGGTT 1200
QY 2397 CCGTGTGGCGCGCTGCCCTCCACCGTTACTGTACTCATGCGCGGTGTGGCGTTTGGT 2456
Db 1201 CCGTGTGGCGCGCTGCCCTCCACCGTTACTGTACTCATGCGCGGTGTGGCGTTTGGT 1260
QY 2457 CACCCACTACTCGAGAGCCCTACCGTTGTCCCACTGCGCCCTTTCGCTGCGAGCGCT 2516
Db 1261 CACCCACTACTCGAGAGCCCTACCGTTGTCCCACTGCGCCCTTTCGCTGCGAGCGCT 1320
QY 2517 GGGCAACTGAGGGGCGCATGAGCGCACCCACAGCGGCTCCCACTCTCTCCCTGCCCAAC 2576
Db 1321 GGGCAACTGAGGGGCGCATGAGCGCACCCACAGCGGCTCCCACTCTCTCCCTGCCCAAC 1380
QY 2577 CTGTGGCTTTGATGTGTGCTGCAAGACCAACCGGGCTCCAGTCCACAGAGAGGA 2636
Db 1381 CTGTGGCTTTGATGTGTGCTGCAAGACCAACCGGGCTCCAGTCCACAGAGAGGA 1440
QY 2637 GGGCAACTGCGCGACCATGAGAAATGCGTGTATCTGCGCAGACTGAGTCTTCATGT 2696
Db 1441 GGGCAACTGCGCGACCATGAGAAATGCGTGTATCTGCGCAGACTGAGTCTTCATGT 1484
QY 2697 GCCACAGGTGTGTCAGTTTCTGCCAGACTGTGGGCGAGTGTGGGGTGAAGGGGAG 2756
Db 1485 ----- 1484
QY 2757 CTGTGTGGAAGTGGATCCGACCACTGCCAGAGTACTGTTCCCTTGGACCTGCGGGG 2816
Db 1485 ----- 1484
QY 2817 CTGTGGACAGAACTGGAGAGGAGTGAAGGACAGAGGCTGGAGGCTGCTGTGGCG 2876
Db 1485 ----- 1484
QY 2877 CTGCATGAGAGAGAGGCTGGAGGGGTTGCCACTGGGGGACCCAGGCGCTGGTGCAA 2936
Db 1485 ----- 1484
QY 2937 AGGCTTGGCTGTAGTTATGTCCTTTGCGCACTCACTACCCACCACTGGCTCGGCA 2996
Db 1485 -----A 1485
QY 2997 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTGCGCGCTCCATAGCGCTCTGC 3056
Db 1486 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTGCGCGCTCCATAGCGCTCTGC 1545
QY 3057 TCATCTGGATAACTGAACAGGACAGCGGCTCCACAGAGAGAAAGCCCTACAAGTG 3116
Db 1546 TCATCTGGATAACTGAACAGGACAGCGGCTCCACAGAGAGAAAGCCCTACAAGTG 1605
QY 3117 CCCCCTCTGTCGTGCTGTGGGCAACCTGGCGCAACTCAAGCGTCACTGTCGATCCA 3176

Db 1606 CCCCCTCTGTCGTGTCGCAACCTGGCCCAACCTCAAGCGTCATGTCGCATCCA 1665
QY 3177 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTTGCAACTACAGCTGCAACAGAGTATGAA 3236
Db 1666 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACAGAGTATGAA 1725
QY 3237 COTCAAAAGTATGCTGCGACACAGGGGAGAGCCCTTCGCTGTGCAACCTGGC 3296
Db 1726 COTCAAAAGTATGCTGCGACACAGGGGAGAGCCCTTCGCTGTGCAACCTGGC 1785
QY 3297 CTATACCAAGGCCACTGGGCAACTACAAAGCTCATCAGAAAGTGCATGCGCATGTTGG 3356
Db 1786 CTATACCAAGGCCACTGGGCAACTACAAAGCTCATCAGAAAGTGCATGCGCATGTTGG 1845
QY 3357 AGCAGGAGGCGCTGCTCTCTGCGCCCTGAGGCTGGGCGCCCACTCATAGCCACCCCTC 3416
Db 1846 AGCAGGAGGCGCTGCTCTCTGCGCCCTGAGGCTGGGCGCCCACTCATAGCCACCCCTC 1905
QY 3417 TGTCTTGGACACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTTCATC 3476
Db 1906 TGTCTTGGACACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTTCATC 1965
QY 3477 AGACTCACCTTGAACCTAAGTGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 3536
Db 1966 AGACTCACCTTGAACCTAAGTGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 2025
QY 3537 TCTCTGATTTTATACAAATGAACCTAGAAACCACTTTTCCCTTTTCCCTGGCTGTCAGG 3596
Db 2026 TCTCTGATTTTATACAAATGAACCTAGAAACCACTTTTCCCTTTCTCCCTGGCTGTCAGG 2085
QY 3597 GGTCCACACAGACTAAGCTAGGCACTATATGGACAGCCCTGAATCCCATGTCAGGGGG 3656
Db 2086 GGTCCACACAGACTAAGCTAGGCACTATATGGACAGCCCTGAATCCCATGTCAGGGGG 2145
QY 3657 CCATATAGACAGGGGACTTGTCTTAGCTCAGTACCAGATGAGCTAAGTATGATAGGCC 3716
Db 2146 CCATATAGACAGGGGACTTGTCTTAGCTCAGTACCAGATGAGCTAAGTATGATAGGCC 2205
QY 3717 TTGATATACCGCCACTGCTCCAGAGGCTATGATGAATGAACTGGTGGAGCTGCCAGCC 3776
Db 2206 TTGATATACCGCCACTGCTCCAGAGGCTATGATGAATGAACTGGTGGAGCTGCCAGCC 2265
QY 3777 TTTTACTGTTTAACTTATTTTCAAGTCTTTTAAATAAAGAAACACTAACAGAAAAAAA 3836
Db 2266 TTTTACTGTTTAACTTATTTTCAAGTCTTTTAAATAAAGAAACACTAACAGAAAAAAA 2325
QY 3837 AAAAAAAAAA 3847
Db 2326 AAAAAAAAAA 2336

RESULT 7
BC058976
LOCUS
DEFINITION BC058976 2220 bp mRNA linear ROD 21-OCT-2003
IMAGE:5685033, complete cds.
ACCSSION BC058976
VERSION BC058976.1 GI:37589246
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryaota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2220)
REFERENCE
AUTHORS
Srausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeng, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F.,
Datchenko, M., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carminci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullihy, S.J., Bobak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalhus, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22386257

12477932

2 (bases 1 to 2220)

Direct Submission

Strausberg, R.

Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letitia Hsiao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zyduerduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 124 Row: a Column: 14

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction.

Location/Qualifiers

1. .2220

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="MGC:67303 IMAGE:5685033"

/tissue_type="Brain, 18.5 dpc, 2-3KB"

/clone_lib="NIH BMAP_EGO"

/lab_host="DH10B"

/notes="vector: pTX-ASC"

1. .2220

/gene="D433028M17Rik"

/db_xref="LocusID:101023"

/db_xref="WGI:2141255"

379..1818

/codon_start=1

/product="D433028M17Rik protein"

/protein_id="AAH58976.1"

/db_xref="GI:37589247"

/db_xref="LocusID:101023"

/translation="MGPERSEGDSSQAGPLPVGLSDSGGGRALSASEVEEPAR
GPEARERGPALCGPGCGAGGPGPLPRLIYSLRLCAFSVSHSSH
LKHQMHSSEKPRCEPCYASQNLNLTHTGKPYRCPHCPACSLGNLR
RQKHTGPTPTCPCTGCCACAPRPPSPTEQSGTMRPSRDLILPLSLHVP
GGASFLPDQOLRGESLCTGSEPLFLLFWTCRGCGELEGEGRLGAMCGR
CMRGEAGVATGPGQGDGFACSLCFPATPHNLHMKTHSEKPRCARCPYA
SAHLNLRKHQRVHTGKPYKPCPLCPYACGNLANLKHGRHISGDKPFRCSLCNYSN

misc_feature

ORIGIN

Query Match 49.9%; Score 1918.8; DB 10; Length 2220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1917 AGTGGATCTGAAGATTCTTTCGACGAAGTCTCTGGGGCCCTCGTGTGTGAGAGCGATT 1976
246 AGTGGATCTGAAGATTCTTTCGACGAAGTCTCTGGGGCCCTCGTGTGTGAGAGCGATT 305

1977 GCTACTAGGCAAGATCTGAGTTTGAAGAGGAGGAGAGGATGATGATGATGATGATGAT 2036
306 GCTACTAGGCAAGATCTGAGTTTGAAGAGGAGGAGGAGGATGATGATGATGATGATGAT 365

2037 CAACGACAGCTCATGGGCTTTGAGAGAGACTCTGAAGGAGAGACTCTCAGGGGGCCAGACC 2096
366 CAACGACAGCTCATGGGCTTTGAGAGAGACTCTGAAGGAGAGACTCTCAGGGGGCCAGACC 425

2097 TGGACTTCCCTATGGGCTGAGTGAAGAGAGACTCTGGGGGGCGGCGGCACATGATGCGGA 2156
426 TGGACTTCCCTATGGGCTGAGTGAAGAGAGACTCTGGGGGGCGGCGGCACATGATGCGGA 485

2157 GAGTGAAGTGTAGGAAACAGCCAGGGGTCCAGGGGGCCAGGGGTGAGAGGGCCAGGCC 2216
486 GAGTGAAGTGTAGGAAACAGCCAGGGGTCCAGGGGGCCAGGGGTGAGAGGGCCAGGCC 545

2217 AGCTGTGAGTGTGGGGGGCCGACAGTGAAGGGCCGCTGTGTGGGGCAGAGAGGCG 2276
546 AGCTGTGAGTGTGGGGGGCCGACAGTGAAGGGCCGCTGTGTGTGGGGCAGAGAGGCG 605

2277 GGGTGG 2336
606 GGGTGG 665

2337 GTCCCATCTACTCGAGGCCACTTGAAGCGGCACATGCACACACACAGCGGGGAGAGCGGTT 2396
666 GTCCCATCTACTCGAGGCCACTTGAAGCGGCACATGCACACACACAGCGGGGAGAGCGGTT 725

2397 CGCTGTGGCGCTGCGCCATACGNTCAGCCAGTGTGTGACACCTGACCGCGACATACCCG 2456
726 CGCTGTGGCGCTGCGCCATACGNTCAGCCAGTGTGTGACACCTGACCGCGACATACCCG 785

2457 CACCCATCTACTGGCGAAGAGCCCTACCGTTGTCCCGCTTTCGCTTCAGCAGCCT 2516
786 CACCCATCTACTGGCGAAGAGCCCTACCGTTGTCCCGCTTTCGCTTCAGCAGCCT 845

2517 GGGCAACCTGAGGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTCTGCGCAAC 2576
846 GGGCAACCTGAGGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTCTGCGCAAC 905

2577 CTGTGGCTTTGATGCTGTGTCCACGACCAACCGGGCTCCCACTCCACAGAGCAGGA 2636
906 CTGTGGCTTTGATGCTGTGTCCACGACCAACCGGGCTCCCACTCCACAGAGCAGGA 965

2637 GGGGACAAATGCCCGGACGATCAGAAAATGCGGTGATCTCTGCGAGACTTGAAGTCTTATGT 2696
966 GGGGACAAATGCCCGGACGATCAGAAAATGCGGTGATCTCTGCGAGACTTGAAGTCTTATGT 1025

2697 GCACACAGTGTGCGCAGCTTTCCTGCGAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGAG 2756
1026 GCACACAGTGTGCGCAGCTTTCCTGCGAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGAG 1085

2757 CTTGTGTGAACTGGATCCGACCACTGCGCAGAGTACTGTTCCTTGGACCTGCGGGG 2816
1086 CTTGTGTGAACTGGATCCGACCACTGCGCAGAGTACTGTTCCTTGGACCTGCGGGG 1145

2817 CTGTGACAGGAACTGGAGAGGGGTGAGGGGAGAGGCTGGGAGCTGCCATGTGTGGCG 2876

Db 1805 TAAGACACAGCCCAACCCCATGTGGCGGGGGGCCCATATGACCAGGAGCACTTG 1864

QY 3683 GCTACGCTACCAGATGAGCTTAAGTATAGGCGCTTGGATTACGCCCACTGCTCCGAGA 3742

Db 1865 ACTGAGGCACTTCATGAGCTCAGTGAGAAGGGCCCTGTATTCACTCCACTGCCCCCAGG 1924

QY 3743 GCCTATGATGAACCTGCTGGGAG-CTGCCACGCTTTTAC-----TGTTTAACTATT 3799

Db 1925 GCGCTGTGACAAACCGCTGCGGAGCTGCCACCTCCCACTGTTATTATTAACTATT 1984

QY 3798 CAGTCTCTTTAATAATAAGGAAACACTAAC 3826

Db 1985 CAGTCTCTTTAATAATAAGGAAACACTAAC 2013

AK096933 2013 bp mRNA linear PRI 15-JUL-2000
 Homo sapiens CDNA FLJ39614 f1s, clone SMINT2000176, weakly similar
 to ZINC FINGER PROTEIN 83.

ACCESSION AK096933
 VERSION AK096933.1 GI:21756542
 KEYWORDS oligo capping, f1s (full insert sequence).

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Mishikawa, I.,
 Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2013)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1. 2013
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SMINT2000176"
 /tissue_type="small intestine"
 /clone_lib="SMINT2"
 /note="cloning vector: pWE18SF13"
 267..1706
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC04903.1"
 /db_xref="GI:21756543"
 /translation="MGFERDSEGLARPGI.PYGLSDSEGGGRALSAEVEEPEV
 GPGEAKERPGACQICGGPTGEGCCGNGGPGGLLPPELLVSCRLCTFVSHY
 LKRMHTSGEKPFRCGRCPYASQLVLRTHRTHGKPYRCFPHCPACCSIGNI
 RHQTHTAGPTTPTCPFCRCPARPSPPEQEGAVPRPDPDALLPLSLHIVH
 GSAFLPDCQLAGEGELCGTGSFPELLFPWTCRCGQGELEGGESILGAAVC

FEATURES
 source

CDS

ORIGIN		Query Match	39.5%;	Score 1518;	DB 9;	Length 1013;
		Best Local Similarity	88.6%;	Pred. No. 0;		
		Matches 1727;	Conservative	0;	Mismatches 178;	Indels 44; Gaps 6;
QY	1885	GAGGAGAGATTCACAGACACCTCCCTCTTTAGTGGATCTGAAGATTCCTTCGACGAA	1944			
DB	102	GAGGAGAGATCCACAGCCCCCTCCCTCTTCAGTGGATCTGAAGATCTCCCTCGACGAA	161			
QY	1945	GGTCTGGGGCCCTGTGTGTGAGAGCGATTTGCTACTAGGCCAAGATCTGGAGTTTCAA	2004			
DB	162	GGACCCGGGGCCCTGTGTATTTAGAGAGTGAATTTGCTACTAGGCCACGATCTGGAGTTTGA	221			
QY	2005	GAGGAGAGAGAGAGATGAAGTGTAGCGGCCACACGACCACTCATGGGCTTTGAGAGA	2064			
DB	222	GAGGAGAGAGAGAGAGAGAGAGGCGACGCGACAGTGCACCTCATGGGCTTCGAGAGA	281			
QY	2065	GACTCTGAAGAGACTCTCAGGGGGCGACACCTGTGACTTCCCTATGGGCTGAGTGCAC	2124			
DB	282	GACTTCGGAAGGACTCTCTGGGGCGACGCTTGGGCTTCCCTATGGGCTGAGCGACGAT	341			
QY	2125	GAGTCTGGGGGGCGCGCCACTTAAGTGGCGGAGAGTGAAGTTGAGGAACCAACCGAGGGT	2184			
DB	342	GAGTCTGGGGGGCGCGCGGCACTTAAGTGGCGGAGAGTGAAGTTGAGAGCCAGCCAGGGT	401			
QY	2185	CCAGGGGAGGCCAGGGGCTGAGAGGCCAGGCCACAGCCCTGTCACTGTGTGGGGGGCCG	2244			
DB	402	CCAGGGGAGGCCAGGGGCTGAGAGGGCCAGGCCACAGCCCTGCACGCTGTGTGGGGGGCCG	461			
QY	2245	GGTGAAGGGCCGTGTTGTGGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCCACGG	2304			
DB	462	GGTGAAGGGCCGTGTTGTGGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCCACGG	521			
QY	2305	TTACTGTACTCATGCGCGCTGTGCGCTTTCGTGTCCCACTACTCTCGAGCCACCTGGAAGCG	2364			
DB	522	CTACTGTACTCATGCGCGCTCTGCACTTCGTGCCACTACTCTCGAGCCACCTGAAGCG	581			
QY	2365	CACATGAGACACACAGCGGGGAGAGCCGTTCCGCTGTGGCGCGTGCACATACGNTCA	2424			
DB	582	CACATGAGACACACAGCGGAGAGAGCCGTTCCGCTGTGGCGCGTGCACATACGCTCA	641			
QY	2425	GCCAGTTCTCAACTGACGACATACCCGACACCACTATCGCGAGAGAGCCCTACCGT	2484			
DB	642	GCCAGTCTGTCAACTGACACACATACCCGACACCACTATCGCGAGAGCCCTACCG	701			
QY	2485	TGTCGCCACTGCCCTTTGCTTGACAGAGCCTGGGCAACCTGAGGGGGAATCAGCGCAC	2544			
DB	702	TGTCGCCACTGCCCTTTGCTTGACAGAGCCTGGGCAACCTGAGGGGGAATCAGCGTACC	761			
QY	2545	CACAGGGGCTCCACTCTCCCTGCCCAACTGTGGCTTTCTGATGCTGTGCTCCACGA	2604			
DB	762	CACAGGAGGGCCCCCACTCTCTCCCTGCCCACTGTGGCTTCCGCTGTCTCTCCACGA	821			
QY	2605	CCAAACCGGGCTCCCACTGCCACAGAGCAGGAGGGGCAATGCCCCGACGATCAGAAAT	2664			
DB	822	CCAGCCGGGCTCCCACTGCCACAGAGCAGGAGGGGGCGGTGCCCGGGGACCTGAAGAT	881			
QY	2665	GGCTGTATCTGCAGACTTGAGTCTTCAATGTGCCACAGGTGGTGGCTTCTCTGCCA	2724			
DB	882	GCTGTGCTCTTCCAGATTTAGAGCTCAATGTGCCACAGGTGGTGGCTTCTCTGCCA	941			
QY	2725	GACTGTGGGCACTCGCGGGTGAAGGGGAGAGCTTGTGTGAACTGGATCCGAACCACTG	2784			
DB	942	GACTGTGGGCACTCGCGGGTGAAGGGGAGGGCCCTCTCGGGCACTGATCAGTAACCACTG	1001			
QY	2785	CCAGAGCTACTGTTCCTTTGACCTGCCGGGCTGTGGAACGGAATCTGAGAGGGGTGAG	2844			
DB	1002	CCAGAGCTACTGTTCCTTTGACCTGCCGGGCTGTGGAACGGAATCTGAGAGGGGTGAG	1061			

2845	QY	GGCAGAGCTGGGAGCTGCATGTGTGGCGCTGCATGGGAGGAGGCTGAGGGGTT	2900
1062	DB	GGTAGTCGGCTGGGAGCTGCCATGTGTGGCGCTGCATGGGAGGAGGCTGAGGGGGT	1121
2905	QY	GCACCTGGGGGAGCCCCAGAGGCGCTGTGTGACAAAGGCTTTGCCCTGTAGTTTATGCCCCCTTT	2964
1122	DB	GCCAGTGGGGGCCCCAGGGCCCCAGTGAACAAGGCTTTGCCCTGTAGCTCTGCCCTTT	1181
2965	QY	GCACCTCACTACCCCAACACCTGGCTCGGCACATGAAGACTCACAGTGGTGAGAAACCC	3024
1182	DB	GCACCTCACTATCCCAACACCTGGCCCGGCACATGAAGACACACAGTGGTGAGAAAGCCC	1241
3025	QY	TTCCGCTGTGCCCGCTGTCCATACAGCGCTCTGTCTCATCTGGATTAACCTGAAACGGCACCCAG	3084
1242	DB	TTCCGCTGTGCCCGCTGTCTTATGCTCTGTCTCATCTGGATTAACCTGAAACGGCACCCAG	1301
3085	QY	CGGCTGCACACAGGAGAAAAGCCCTCAAGTGGCCCTCTGTCCGATGATGCTGTGGCAAC	3144
1302	DB	CGGCTGCATACAGGAGAGAGCCCTCAAGTGGCCCTCTGTCCCTTATGCTGTGGCAAT	1361
3145	QY	CTGGCCAACTCAAGGCTCATGTGCGCATCCACTCTGGTGACAAACCTTTTCGTGTAGC	3204
1362	DB	CTGGCCAACTCAAGGCTCATGTGCGCATCCACTCTGGTGACAAACCTTTTCGTGTAGC	1421
3205	QY	CTTTGCAACTACAGCTGCAACCGAGATATGAACTTCAAACGTCTATGCTGCGACACAGC	3264
1422	DB	CTTTTGAACCTTACAGCTGCAACCGAGAGCATGAACCTTCAAACGTCTATGCTGCGGACACA	1481
3265	QY	GGGAGAAAGCCCTTCGCTGTGCCACTGGCCCTTATACACAGGCCACTTGGGACAACTTAC	3324
1482	DB	GGGAGAAAGCCCTTCGCTGTGCCACTGGCCCTTATACACAGGCCACTTGGGACAACTTAC	1541
3325	QY	AAGGCTCATCAGAAAGTGCATGGCCATGTGTGAGCAGAGGGCCCTGTCTCTGCGCCCT	3384
1542	DB	AAGCGCCACCAGAAAGTGTATGGCCACGGTGGGGCAGGAGGGCTGTGTCTCTGCGCTCT	1601
3385	QY	GAGGGCTGGGCCCCCACCCTCATAGCCCCACCTCTGTCTTTGAGCACTCGGGGTCCAGCAGCC	3444
1602	DB	GAGGGCTGGGCCCCCACCCTCATAGCCCCACCTCTGTCTTTGAGCTCTCGGGGCCACCGAGCC	1661
3445	QY	CTGGGTCCTACTGTGTAGAGGGCTCTTCATTACAGACTCACCTTGAACCTAAGTTCCTT	3504
1662	DB	CTGGGAGACTGTGTGACCGGGCTGTCCACAGACTCATCTCTGAAC-----	1708
3505	QY	TTACCTGGGGCTCTAGAAATAGCCCTATGCTCTGCTCATTTTATACAAATGAACCTAGAA	3564
1709	DB	-----TAGTCTCTTCTCCCATGTTTTATACAGACGGACCGAGAA	1748
3565	QY	ACCACTTTCCCTTTCTCCCGCTGTGTGAGGGCTCCACACAGACTTAACCTTAGGCCTA	3624
1749	DB	GCCACCTT----TTTCTCCCGCTGCCAGGGGCTCCACACAGACTTAACCTTAGGCCTA	1804
3625	QY	TATGGACAGGCTCAATCCCATGTGTAGGGGG--CCATATAGACACAGGGGA--CTTGCTCTTA	3684
1805	DB	TAGGACAGGCCCAACCCCATGTGGCGGGGGGCCCATATGACACAGGGGACCTTGCCCTTG	1864
3683	QY	GCTCACTGTAACAGATGAGCTTAAGTGTATGGGCGCTTGGATTCAAGGCACCTGTCTCCAG	3742
1865	DB	ACTGAGGCACCTTTCATGAGCTCAGTGTAGAAGGGCGCTGTATTCACTCCACTGCCCGCCAGG	1922
3743	QY	GGCTATGATGAACCTGGTTTGGAG--CTGCCACAGCCTTTTAC-----TGTTTTAACTTATTT	3799
1925	DB	GGCTGTGGAACAAACCGGCTGTGGGAGTCTCCACAGCTCCCACTGTATTATTAACCTTATTT	1981
3798	QY	CAGTCTTTTATAATAAGAAACACTAATAC	3826
1985	DB	CAGTGTCTTTATAATAAGAAACACTAATAC	2013

RESULT 10	
BC052282	
LOCUS	
BC052282	2159 bp
mRNA	linear
PRI 09-OCT-2000	

854	DB	GGGCAACTGAGCGGCATCAGGTGATCCACGACGAGGGCCCCCACTCTCTCCCTGCCCCGAC	913
2577	QY	CTGTGGCTTTTCGATGTGTGCTCCACGACCAACCCGGGCTCCGAGTCCACAGACAGGA	2636
914	DB	CTGTGGCTTCGGCTGCTACTCCACGACGAGCCGGCTCCAGTCCACAGACAGGA	973
2637	QY	GGGGCAAATGCCCGGACGATCAGAAATGGCTGATCTCTGCCAGACTTGAGTCTTCATCT	2696
974	DB	GGGGGGGGTGCCTCCCGCGACCTGAAGATGCTCTGCTCTTCCAGATTGAGCCTCCATGT	1033
2697	QY	GCCACAGGTGGTGCACGATTTCTGCCAGACTGTGGCGAGCTCGGGGGTGAAGGGGAGAG	2756
1034	DB	GCCACAGGTGGTGCCAGTTTCTTGCCAGACTGTGGCGAGCTCGGGGGTGAAGGGGAGGG	1093
2757	QY	CTTGTGTGAACTGGATTCGAAACCACTGCCAGAGCTACTGTTCCTTTGACCTGCCGGG	2816
1094	DB	CCTCTCGGGACTGGATTCAGAAACCACTGCCAGAGCTGCTATTCCCTTGGACCTGCCGGGG	1153
2817	QY	CTGTGACAGAACTGGAGAGGGTGAAGGCACAGAGCTGGGAGCTGCATGTGTGGGGG	2876
1154	DB	CTGTGACAGAGCTGGAGAGGGTGAAGGGTGAAGGGTGAAGGGTGAAGGGTGAAGGGT	1213
2877	QY	CTGCATGCGAGGAGAGGCTGGAGGGTGGCACTGGGGGACCCAGGGCCCTGGTGACAA	2936
1214	DB	CTGCATGCGAGGAGAGGCTGGAGGGTGGCACTGGGGGACCCAGGGCCCTGGTGACAA	1273
2937	QY	AGGCTTGGCTGTAGTTTATGCCCTTTGCGCACTCACTACCCCAACCACTGGCTCGGCA	2996
1274	DB	AGGCTTGGCTGTAGGCTCTGGCCCTTTGGCCATCACTATCCCAACCACTGGCCCGGCA	1333
2997	QY	CATGAAGACTCAAGTGGTGAGAAACCTTTCCGCTGTGCCCGTGTCCATPAGCCTCTGC	3056
1334	DB	CATGAAGACACAGTGTGAGAAAGCCCTTCGCTGGCCCGCTGTCTTATGCTCTCTGC	1393
3057	QY	TCATCTGGATTAACCTGAAACCGCACAGCGGCTGCCACAGAGGAAAGCCCTACAGTG	3116
1394	DB	TCATCTGGATTAACCTGAAACCGCACAGCGGCTGCCATACAGAGGAAAGCCCTACAGTG	1453
3117	QY	CCCCCTCTGTCGTATGCCCTGTGCACACTGGCCAACTCAAGCGTCATGGTGGCATCCA	3176
1454	DB	CCCCCTCTGCCCTTATGCTGTGGCACTGTGGCCAACTCAAGCGTCATGGTGGCATCCA	1513
3177	QY	CTCTGTGACAAACCTTTTCGGTGTAGCTTTGGCACTACAGCTGMAACAGAGTATGAA	3236
1514	DB	CTCTGTGACAAACCTTTTCGGTGTAGCTTTGGCACTACAGCTGMAACAGAGTATGAA	1573
3237	QY	CCTCAACGTCATAGCTGTGGACACACGCGGAGAGGCCCTTCGGTGTGCCACCTGCGC	3296
1574	DB	CCTCAACGTCATAGCTGTGGACACACGCGGAGAGGCCCTTCGGTGTGCCACCTGCGC	1633
3297	QY	CTATACCAAGGCCACTGGGACAACTACAGCGTCATCAGAAGGTGCATGGCCATGGTGG	3356
1634	DB	CTATACCAAGGCCACTGGGACAACTACAGCGTCATCAGAAGGTGCATGGCCATGGTGG	1693
3357	QY	AGCAGGAGGCTGTGTCTCTGCCCCGTGAGGCTGGCCCCCACTCATAGCCCACTCTC	3416
1694	DB	GGCAGGAGGCTGTGTCTCTGCCCCGTGAGGCTGGCCCCCACTCATAGCCCACTCTC	1753
3417	QY	TGTTTTGAGCACTCGGGGTCCAGCAGGCCCTGGGTGTACTGTGTAGCAGGCTCTTCAATC	3476
1754	DB	TGTTTTGAGCTCTCGGGGCCCAACGACCTTGGGACTGTCTGGCAGCGGCTGTGCACAC	1813
3477	QY	AGACTCACTTGAACCTAAGTTCTTTTACCTGGGGCTCTAGGAATTAGCCTATGCC	3536
1814	DB	AGACTCACTTGAACCTAAGTTCTTTTACCTGGGGCTCTAGGAATTAGCCTATGCC	1840
3537	QY	TCCTGCAATTTATACAAATGAACTAGAAAACACCTTTCCCTTTCTCCCCGGCTGGTCAGG	3596
1841	DB	CCCCATGTTTTATACAGCGGACCAAGGACCACTT-----TTTCTCCCCGTGGCCAGG	1896
3597	QY	GGCTCCACACAGACTAACCTTAGGCATCTATATGACACAGCTGTAATCCATGTGTGAGGGG	3656
1897	DB	GGCTCCACACAGACTAACCTTAGGCATCTATATGACACAGCTGTAATCCATGTGTGAGGGG	1956

RESULT	11	AX274856	1975 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	AX274856					
DEFINITION	Sequence	121 from Patent WO0172777.				
ACCESSION	AX274856					
VERSION	AX274856.1	GI:16547488				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Hillman J.L., Baughn M.R., Yue H., Lal P., Lu D.A., Patterson C.,					
	Animal Y., Bandman O., Tang Y.T., Mathur P., Shah P., Au-Young J,					
	and Reddy R.					
TITLE	Transcription factors					
JOURNAL	Patent: WO 017277-A 121 04-OCT-2001;					
	Incyte Genomics, Inc. (US)					
FEATURES	Location/Qualifiers					
source	1..1975					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	/note="Incyte ID No: 870100CB1"					
ORIGIN						
Query Match	38.9%;	Score 1498;	DB 6;	Length 1975;		
Best Local Similarity	88.6%;	Pred. No. 0;				
Matches 1707;	Conservative	0;	Mismatches 176;	Indels 44;	Gaps	6;
QY	1917	AGTGGATCTGAAGATTCCTTCGACGAAGTCCTGGGGCCCTGGTGTGGAGAGCGATTT	1976			
Db	86	AGTGGATCTGAAGATTCCTTCGACGAGAACCCGGGGCCCTGGTATTTGGAGAGTGATTT	145			
QY	1977	GCTACTAGGCCAAGATCTGGAGTTTGAAGAGGAAGAGAGAGATGAAGGTGACGGCCA	2036			
Db	146	GCTACTAGGCCAGGATCTGGAGTTTGAAGAGGAAGAGAGAGAGAGAGCGACGGCAA	205			
QY	2037	CAACGACCACTCATGGGCTTTCGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC	2096			
Db	206	CAGTGACCACTCATGGGCTTCGAGAGAGACTCGAAGGAGACTCTCTGGGGGCCAGGCC	265			
QY	2097	TGGACTTCCCTATGGGTGAGTACGACGAGTCTGGGGGGGGCCGCCACTAAGTGGCGA	2156			
Db	266	TGGCTTCCCTATGGGCTGAGCGACGATGATCTGGGGGGGGCCGCCACTAAGTGGCGA	325			
QY	2157	GAGTGAAGTTGAGGAACACGCCAGGGGTTCAGGGAGGCCAGGGGTGAGAGGCCAGGCC	2216			
Db	326	GAGTGAAGTTGAGGAGCAGCCAGGGGTTCAGGGAGGCCAGGGGTGAGAGGCCAGGCC	385			
QY	2217	AGCCTGTCAGCTGTGTGGGGGGCCGACAGGTGAGGGGCCGCTGTTGTGGGGCAGGAGGGCG	2276			
Db	386	AGCCTGCCACTGTGTGGGGGGCCGACAGGTGAGGGGCCGCTGTTGTGGGGCAGGAGGGCC	445			
QY	2277	GGGTGGGGGGCCCCCGCTGCCCCCAACGGTTACTGTACTCATGCGCGTGTGCGCTTCGT	2336			

1526	Db	GCACGAGGGCGCTGCTCTCTGCTCTCAGGGCTGGGCCCCACCTCATAGCCCCCCTC	1586
3417	QY	TCGTTTGGACATCCTGGGGTCCAGACCCCTGGGTGCTACTGGTAGCAGGCTCTTCATTC	3476
1586	Db	TGTTTGTAGCTCTCGGGGCCACACGAGCCCTGGGGACTGCTGGCAGACGGGCTGTCACAC	1645
3477	QY	AGACTCACCCTGAACTAACTAGTTCCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC	3536
1646	Db	AGACTCATCCTGAAC-----TAGTCTCTTCTT	1672
3537	QY	TCCTGCATTTTATACAAATGAAGTAAACCCACCTTTCCCTTTCTCCCGCGTGGTCAGG	3596
1673	Db	CCCCATGTTTTATACAGCGGACCAAGACCACCTT-----TTTCTCCCGCTGGCCAGG	1728
3597	QY	GGCTCCACACAGACTAACTAGGCACTATATGACAGCCTGNAATCCATGCTCAGGGGG	3656
1729	Db	GGCTCCACACAGACTAACTAGGCACTATATGAGACACAGCCCAACCCCATGGCGGGGGGG	1788
3657	QY	-CCATATAGACAGGGGA-CTTGCTCTTAGCTCACGTACCAGATGAGTAACTAGTTAGGG	3714
1789	Db	CCCATATGACACAGGGGACTTGCTCTGACTGAGGCACITTCACAGCTCAGTGAGAAGG	1848
3715	QY	CTTTGAAATCACGGCACCTGCTCCAGAGCTAATGATGAACCTGGTTGGAG-CTGCCCA	3773
1849	Db	CCCTGTATTCACCTCCACTGCCCGCCAGGGGCTGTGACAAACCGGCTGGGGGACTGCCCA	1908
3774	QY	GCTTTTAC-----TGTTTAACTTATTTTTCAGTCTCTTATAATAAGAAACACTAACAGA	3829
1909	Db	GCTCCCACTGTTTATTTAACTTATTTTTCAGTCTCTTATAATAAGAAACACTAACAA	1968
3830	QY	AAAAAA 3836	
1969	Db	AAAAAA 1975	

RESULT 12	1996 bp	DNA	linear	PAT 17-JUL-2003
BD242880	Secreted proteins and polynucleotides encoding them.			
LOCUS	BD242880			
DEFINITION	BD242880			
ACCESSION	BD242880.1 GI:33052650			
VERSION	JP 2002536973-A/31.			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1996)			
AUTHORS	Valenzuela,D., Yuan,O., Hoffman,H., Hall,J. and Rapiejko,P.			
TITLE	Secreted proteins and polynucleotides encoding them			
JOURNAL	Patent: JP 2002536973-A 31 05-NOV-2002;			
COMMENT	OS Homo sapiens (human) PN JP 2002536973-A/31 PD 05-NOV-2002 PF 18-FEB-2000 JP 2000599860 PR 19-FEB-1999 US 60/120680,23-APR-1999 US 09/298733 PR 17-AUG-1999 US 60/149639,23-SEP-1999 US 60/155686 PR 21-OCT-1999 US 60/157247,29-NOV-1999 US 60/167823 PR 09-NOV-1999 US 60/167822,15-FEB-2000 US 60/182711 PI DARIO VALENZUELA,OLIVE YUAN,HEIDI HOFMAN,JERFF HALL,PETER PI RAPIEJKO PC C12N15/09,A61K38/00,A61P3/10,A61P5/14,A61P11/00,A61P11/06,PC A61P19/02, PC A61P21/04,A61P25/14,A61P27/02,A61P29/00,A61P31/04,A61P31/10, PC A61P31/12, PC A61P31/18,A61P31/20,A61P31/22,A61P37/00,A61P37/06,C07K14/435, PC C12N5/10, PC C12P19/34/(C12P19/34,C12R1:91),C12N15/00,C12N5/00,A61K37/02 CC Secreted proteins and polynucleotides encoding them FH Key Location/Qualifiers FT source 1..1996 /organism='Homo sapiens (human)'. FT Location/Qualifiers 1..1996			

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="taxon:9606"
```

ORIGIN

Query Match	35.5%	Score 1367.6	DB 6	Length 1996	
Best Local Similarity	86.2%	Pred. No. 0			
Matches 1598	Conservative	0	Mismatches 210	Indels 45	Gaps 6
QY	2003	AAGAGGAAGAGGAAGAGATGAAGAGTGCACGGCCACACAGCACAGCTCATGGGCTTTGAGA	2062		
DB	139	AAGTGAAGAAGGGAGAGTTTGGAGCCTGATCCTACCCATGCTGATGCTCTCTCTTATGT	198		
QY	2063	GAGACTCTGAAGAGACTCTCAGGGGCCAGACCTGSACTTCCCTATGGGCTGAAGTGAAG	2122		
DB	199	CTATTTTACCAGGAGACTCTCTGGGGGCCAGGGCTTGGCTTCCCTATAGGGCTCAGCGAGC	258		
QY	2123	ACAGAGTCTGGGGGGCGGCGCGCACTAAGTGCAGAGAGTGAAGTTGAGGAACCAAGCCAGGG	2182		
DB	259	ATGAGTCTGGGGGGCGGCGCGCACTAAGTGCAGAGAGTGAAGTTGAGGAGCCAGCCAGGG	318		
QY	2183	GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCGTCTCAGCTGTGTGGGGGGCCGA	2242		
DB	319	GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCGTCTCAGCTGTGTGGGGGGCCGA	378		
QY	2243	CAGGTGAGGGGGCGGTGTGTGTGGGCGAGGAGGGGGTGGGGGGCCCCCGCTGCCCCCAC	2302		
DB	379	CAGGTGAGGGGGCGGTGTGTGTGGGCGAGGAGGGGCCGGTGGGGGGCCCCCTGTGCCCCCAC	438		
QY	2303	GGTTACTGTACTCATGCCGCTGTGCCCTTTCGTGCCCATCTACTCAGAGCCACTGAGC	2362		
DB	439	GGCTACTGTACTCATGCCGCTTTCGACCTTCGTGTCCCACTACTCAGAGCCACTGAGC	498		
QY	2363	GGCACATGCGACACACAGCGGGAGAACGGTTCCGTGTGGCCGTGCCCATACGCNT	2422		
DB	499	GGCACATGCGACACACAGCGGGAGAACGGTTCCGTGTGGCCGTGCCCATACGCCT	558		
QY	2423	CAGCCAGTTTCGTCACCTCAGCGGACATACCCGACCCACTACTGCGGAGAGCCCTACC	2482		
DB	559	CAGCCAGCTCGTCACCTCAGCGGACATACCCGACCCACTACTGCGGAGAGCCCTACC	618		
QY	2483	GTTGTCCCCACTGCCCTTTGCCCTGCGAGCAGCTTGGGGAACCTCAGCGCGCATCAGCGCA	2542		
DB	619	GCTGTCCCCACTGTGCCCTTTGCCCTGCGAGCAGCTTGGGCAACCTCAGCGCGCATCAGCGTA	678		
QY	2543	CCACACAGGGGCTCCCACTCTCCCTGCCCAACCTGTGGGCTTTTCGATGCTGTGCTCCAC	2602		
DB	679	CCACAGCAGGGCCCCCACTCTCTCCCTGCCGACCTGTGGGCTTCGGCTGTGTACTCCAC	738		
QY	2603	GACCAAACCGGCTCCGAGTCCCAAGAGCAGGAGGGGAACATGCCCCGACATCAGAA	2662		
DB	739	GACCAAGCCCGGCTCCAGTCCCAAGAGCAGGAGGGGGCGGTGCCCGGCGACTGAAG	798		
QY	2663	ATGGCTGTATCTCTGCCAGACTTGAAGTCTTCATGTGCCACAGGTGTGTCAGTTTCTCTGC	2722		
DB	799	ATGCTCTGCTCTCTCCAGATTGAGCCTCATGTGCCACAGGTGTGTGCCAGTTTCTCTGC	858		
QY	2723	CAGACTGTGGGACGCTCGGGGTGAAGGGGAGAGCTGTGTGGAACATGATTCGAAACCAAC	2782		
DB	859	CAGACTGTGGGACGCTCGGGGTGAAGGGGAGGGCCCTCTCGGGAGCTGGATCAGAAACCAAC	918		
QY	2783	TGCCAGAGCTACTGTCTCCCTTGGACCTTCCGGGGCTGTGGACACAGAACTGGAGAGGGTG	2842		
DB	919	TGCCAGAGCTGCTATTCTCTTGGACCTTCCGGGGCTGTGGACAGAGCTTGGAGGGGTG	978		
QY	2843	AGGCGAGAGGCTGGAGCTGCCATGTGTGGGCGCTGCATCGAGAGAGAGGCTGGAGGGG	2902		
DB	979	AGGCTAGTCGGCTGGGAGCTGCCATGTGTGGGCGCTGCATCGAGAGAGAGGCTGGAGGGG	1038		
QY	2903	TTGGCACTGGGGGAGCCCCAGGGGCCCTGTGTGACAAAGGCTTTTGCCTGTAGTTTATGCCCT	2962		
DB	1039	GTGCCAGTGGGGGGCCCCAGGGGCCCTGTGTGACAAAGGCTTTTGCCTGTAGCTCTGCCCT	1098		

[illegible]

RESULT 13
AC074117
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCES

AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 195880)
 AUTHORS Sun, H., Kozlowicz, A. and Dignan, G.
 TITLE The sequence of Homo sapiens BAC clone RP11-413M20
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 6 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 7 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Feb 1, 2002 this sequence version replaced gi:14495402.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: H_NH0413M20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-413M20 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Cacanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong

and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-413M20; actual end is at base position 195880 of RP11-413M20.

A transposon exists in the growth of this clone which has been deleted from the finished sequence. This transposon would insert after base 78210. Unsure bases exist between bases 165498 to 165505. Unresolved tandem repeat exists between bases 4700 to 6435. Polymorphisms have been identified between AC009427 and AC074117. Data from AC009427 was used to finish this clone, AC074117.

FEATURES

	Location/Qualifiers
source	1..195880 /organism="Homo sapiens" /mol_type="Genomic DNA" /db_xref="taxon:9606" /chromosome="2" /map="2" /clone="RP11-413M20" /clone_lib="RP11-413M20"
repeat_region	1..207 /rpt_family="L1"
repeat_region	208..500 /rpt_family="Alu"
repeat_region	501..641 /rpt_family="L1"
repeat_region	746..1042 /rpt_family="Alu"
repeat_region	1221..1538 /rpt_family="Alu"
repeat_region	1518..1538 /rpt_family="(A)n"
repeat_region	1546..1866 /rpt_family="L1"
misc_feature	2505..2520 /note="Similar to Homo sapiens EST AW327943 (NID:G6798438)"
misc_feature	2507..2520 /note="Similar to Homo sapiens EST AW327944 (NID:G6798439)"
repeat_region	2722..2766 /rpt_family="AT-rich"
repeat_region	2959..3251 /rpt_family="Alu"
misc_feature	3153..3202 /note="Similar to Homo sapiens EST AU130518 (NID:G10990872)"
repeat_region	3558..3645 /rpt_family="MER1_type"
repeat_region	3646..3952 /rpt_family="Alu"
repeat_region	3651..3671 /rpt_family="(TTTTTA)n"
repeat_region	3953..3995 /rpt_family="MER1_type"
repeat_region	3962..3990 /rpt_family="AT-rich"
repeat_region	3996..4284 /rpt_family="Alu"
repeat_region	4285..4331 /rpt_family="MER1_type"
repeat_region	4515..4593 /rpt_family="L2"
repeat_region	4826..4870 /rpt_family="(TCTCCC)n"
misc_feature	4780..6423 /note="CpG island (GC=74.2, o/e=0.70, #CpGs=131)"
repeat_region	5543..5910 /rpt_family="SVA"
repeat_region	5926..5950

Db 168382 CCACGGTGGGAGAGAGGCTGTCTCTCTGCTCTGAGGGCTGGGCCCCACCTCATAG 168441
 Qy 3408 CCCACCTCTCTGTTTGTAGACATCGGGGTCAGAGCCCTGGGTGCTACTGTGTAGCAGGC 3467
 Db 168442 CCCACCTCTCTGTTTGTAGCTCTGGGGGCCACAGCCCTGGGAGCTGTGGCAGCCGGC 168501
 Qy 3468 TCTTCATTACAGCTCACCTTGAACCTAGTCTTTTACCTGGGGCTCTAGGAATTAG 3527
 Db 168502 TGTCCACACAGACTCATCTGAC-----TAG 168528
 Qy 3528 CCCTATGCTCTCTGATTTATPACAAATGAATAGAAACCACTTTCCCTTTCTCCCGC 3587
 Db 168529 GTCTCTCTCTCCCATGTTTATACAGCGGACCAAGCCACCTT---TTCTCTCCCGC 168584
 Qy 3588 CTGGTACGGGCTCCACAGCAGCTACCTAGGCTATATGACAGCCTGAATCCCATG 3647
 Db 168595 CTGGCCAGGGGCTCCACAGACTAACGTAGGCATATAAGACAGGCCCAACCCCATGG 168644
 Qy 3648 GTACAGGGG-CCATATAGACAGGGGA-CTTGTCTTTAGCTCACGTACAGATGAGCTAAG 3705
 Db 168645 GCGGGGGGCCCCATATGACAGGGGACCTTGCTTGACTGAGGCATTTACGAGCTCAG 168704
 Qy 3706 TGATTAGGCTTGTGATTCACCGCCACTGCTCCAGAGGCTATGATGAATGTTGGGA 3765
 Db 168705 TGAGAAGGGCCCTGTATTACCTCCACTGCCGCCAGGGGCTGTGGACAAACCGGCTGGG 168764
 Qy 3766 G-CTGCCAGGCTTTTAC-----TGTTTAACTTATTTTCACTTATTAATAAAGGAAC 3820
 Db 168765 GACTGCCAGCTCCACCTGTTTATTAACTTATTTCACTTATTAATAAAGGAAC 168824
 Qy 3821 ACTAACAGA 3829
 Db 168825 ACTAACAA 168833

RESULT 14
 BC040650
 LOCUS
 DEFINITION Homo sapiens, similar to hypothetical protein FLJ32203, clone IMAGE:479511, mRNA.
 ACCESSION BC040650
 VERSION BC040650.1 GI:26996506
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1461)
 Strausberg, R.
 Direct Submission
 Submitted (29-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Boedet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasia van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 70 Row: d Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: Excellent library of origin.

FEATURES
 source
 Location/Qualifiers
 1..1461
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:479511"
 /issue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC 96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 29.0%; Score 1115.4; DB 9; Length 1461;
 Best Local Similarity 87.2%; Pred. No. 1.2e-257;
 Matches 1307; Conservative 0; Mismatches 147; Indels 44; Gaps 6;
 Qy 2352 CCACCTGAGCGGCACATGACAGACACACAGCGGGAGAGCGGTCCTGCTGGCGCGCTG 2411
 Db 1 CCACCTGAGCGGCACATGACAGACACACAGCGGGAGAGCGGTCCTGCTGGCGCGCTG 60
 Qy 2412 CCATACGCTACAGCGGCTGCTCAACCTGACGCGGACATACCGCACCCATATCTGGCA 2471
 Db 61 CCCTACGCTACAGCGGCTGCTCAACCTGACGCGGACATACCGCACCCATATCTGGCA 120
 Qy 2472 GAAGCCCTACGCTGCTCCCACTGCGGCTTGGCTGACAGAGCTGGCAACCTGAGGG 2531
 Db 121 GAAGCCCTACGCTGCTCCCACTGCGGCTTGGCTGACAGAGCTGGCAACCTGAGGG 180
 Qy 2532 GCATCAGCGCACCCACACAGCGGCTGCTCCCACTGCTCCCTGCCAACCTGTGGCTTGCATG 2591
 Db 181 GCATCAGGTTACCCACGCGAGGGCCCCCACTGCTCCCTGCCGACCTGTGGCTTCCGCTG 240
 Qy 2592 CTGTGCTCCAGACCAACCGCGCTCTCCAGTCCCAAGAGCAGAGAGGGGAGCAATGCCCC 2651
 Db 241 CTGTACTCCAGCACCGCGGCTCCCACTGCTCCCAAGAGCAGAGAGGGGCGGTGCCCC 300
 Qy 2652 ACATCAGAAATGGCGTATCTGCGAGCTTGTAGTCTTATGTCACACAGGCTGTGC 2711
 Db 301 CGGACCTGAGATGCTCTGCTCTTCCAGATTTGAGCTTCCATGTGCCACAGGTGGTGC 360
 Qy 2712 CAGTTTCTCCAGACTGTGGGAGCTGTGGGCTGAGGGGAGAGCTGTGTGGAAGCTGG 2771
 Db 361 CAGTTTCTCCAGACTGTGGGAGCTGTGGGCTGAGGGGAGAGGCTCTGTCCGGGACTGG 420
 Qy 2772 ATCCGACCACTGCCAGACTGTGTTCCCTGACCTGCGGGGCTGTGACAGAGACT 2831
 Db 421 ATCAGAACCACTGCCAGACTGTATTTCCCTTGAGCTGCGGGGCTGTGACAGAGCT 480
 Qy 2832 GGAGGGGCTGAGGGCAGAGGCTGAGGCTGCGAGCTGCGCTGTGGCGCTGTGACAGAGA 2891
 Db 481 GGAGGGGCTGAGGGTACTCGGCTGGAGCTGCGCTGTGGCGCTGTGACAGAGA 540
 Qy 2892 GGCTGAGGGGTTGCCACTGCGGAGCCCCAGGGGCTGTGACAGAGGCTTGCCTGTAG 2951
 Db 541 GGCTGAGGGGTTGCCAGTGGGGGCCCCAGGGGCTGTGACAGAGGCTTGCCTGTAG 600
 Qy 2952 TTTATGCCCTTTGCCACTCACTACCCCAACCACTGGCTCGGCACATGAAGACTCACAG 3011
 Db 601 CCTCTGCCCTTTGCCACTCACTATCCCAACCACTGGCCCGGCACATGAAGACACAG 660
 Qy 3012 TGGTGAAGAACCTTCCGCTGTGCCGCTGTGCATACGCTCTGCTCATCTGATACCT 3071
 Db 661 TGGTGAAGAACCTTCCGCTGTGCCGCTGTGCCTTATGCTCTGCTCATCTGATACCT 720
 Qy 3072 GAACGCGCACCGGCTGCCACACAGGAGAAAGCCCTACAAGTCCCTCTGTCGGTA 3131

Db 721 GAACGGCACCAGCGTCCATACAGGAGAGCCCTACAGTGCCTCTGCGCTTA 780
 Qy 3132 TGCTGTGGCAACCTGGCAACCTCAAGCGTCAATGTGCGATCCACTGTGTGACAAACC 3191
 Db 781 TGCTGTGGCAATCTGGCAACCTCAAGCGTCAATGTGCGATCCACTGTGTGACAAACC 840
 Qy 3192 TTTTTCGGTGTAGCTTTTGGCACTACAGCTGCAACAGAGTATGAACCTCAACAGCTCATAT 3251
 Db 841 TTTTTCGGTGTAGCTTTTGGCACTACAGCTGCAACAGAGTATGAACCTCAACAGCTCATAT 900
 Qy 3252 GCTGCGACACACGGGAGAGCCCTTCGGCTGTGCGACCTGCGCTTATACACAGGCGCA 3311
 Db 901 GCTGCGGCAACACGGGAGAGCCCTTCGGCTGTGCGACCTGCGCTTATACACAGGCGCA 960
 Qy 3312 CTGGGCAACTCAACGGTCAATCAGAGGTGATGCGCATGTGAGAGAGGAGGCGCTGG 3371
 Db 961 CTGGGCAACTCAACGGGCAACAGAGGTGATGCGCATGTGAGAGAGGAGGCGCTGG 1020
 Qy 3372 TCTCTCTGCGCTGAGGGCTGGGCGCCACCTCTATAGCCACCTCTGTTTGTGAGCACTCG 3431
 Db 1021 TCTCTCTGCGCTGAGGGCTGGGCGCCACCTCTATAGCCACCTCTGTTTGTGAGCTCTCG 1080
 Qy 3432 GGTTCAGAGCGCTGGTGTCTACTGTGTAGAGGGCTCTTCAATTCAGACTCACTTGAAC 3491
 Db 1081 GGGGCCACACAGCGCTGGGAGCTGTGCGACCGGGCTGTCCACACAGACTCATCTCTGAAC 1140
 Qy 3492 TAACTAGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCTCTGCAATTTTATAC 3551
 Db 1141 -----TAGTCTCTCTTCCCGATGTTTATAC 1167
 Qy 3552 AATGAACTAGAACACCTTTCCTTCTCCCGCTGTGTAGGGCTTCCACAGACT 3611
 Db 1168 AGCGGACCAAGAGCCACCTT-----TTTCTCCCGCTGTGTAGGGCTTCCACAGACT 1223
 Qy 3612 AACCTAGGCACTATATGAGCAGAGCTCAATCCCATGTGTAGGGGCGC-ATATAGACAGG 3670
 Db 1224 AACTAGGCACTATATGAGCAGAGCAGCCCAACCCATGAGGGGGGGGCCAATATGAGCAGG 1283
 Qy 3671 GGA-CTTGTCTTACTCAGCTACAGTACAGATGAGTAAAGTATGAGGCTTGGATTCACCGC 3729
 Db 1284 GGA-CTTGTCTTACTCAGGCACTTATGAGTACAGTACAGTACAGGAGGAGGCTTATACCTC 1343
 Qy 3730 CACTGCTCCAGAGCTATGAGTAACTGTGTGGAG-CTGCCAGCTTTTAC-----TG 3784
 Db 1344 CACTGCCCCAGGGGCTGTGAGCAACCGCTGGGGAGCTGCCAGCTCCACCTGTTT 1403
 Qy 3785 TTTTAACTTTTCACTGCTTTTATATATAGGAAACACTAACAAAAA 3842
 Db 1404 ATTTAACTTTTCACTGCTTTTATATATAGGAAACACTAACAAAAA 1461

RESULT 15
 AX714442 1456 bp DNA linear PAT 15-APR-2003
 LOCUS
 DEFINITION Sequence 1126 from Patent EP1293569.
 ACCESSION AX714442
 VERSION AX714442.1 GI:29889394
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Igano, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.
 Full-length cDNAs
 Patent: EP 1293569-A 1126 19-MAR-2003;
 JOURNAL Helix Research Institute (JP); Research Association for Biotechnology (JP)
 FEATURES
 Location/Qualifiers
 1..1456

ORIGIN
 Query Match 29.0%; Score 1115.2; DB 6; Length 1456;
 Best Local Similarity 87.4%; Pred. No. 1.3e-257;
 Matches 1305; Conservative 0; Mismatches 144; Indels 44; Gaps 6;
 Qy 2341 CACTACTCGAGCGCACCTGAAAGCGGCATGACAGACACAGCGGGGAGAGCGTTCCTCGC 2400
 Db 1 CACTACTCGAGCGCACCTGAAAGCGGCATGACAGACACAGCGGGGAGAGCGTTCCTCGC 60
 Qy 2401 TGTGGCGCTGCGCCATACGCTCAGCCAGTGTGTCTCAACCTGACGACATACCGCACC 2460
 Db 61 TGTGGCGCTGCGCCATACGCTCAGCCAGTGTGTCTCAACCTGACGACATACCGCACC 120
 Qy 2461 CATCTGCGAGAGAGCGCTACCGTGTGCGCCACTGCGCCCTTGCCTGACAGCGCTGGGC 2520
 Db 121 CACACTGCGAGAGAGCGCTACCGTGTGCGCCACTGCGCCCTTGCCTGACAGCGCTGGGC 180
 Qy 2521 AACCTGAGCGCGCATCAGCGCACCCACAGCGGCTCTCCACTCTCTCTCTGCCAACCTGT 2580
 Db 181 AACCTGAGCGCGCATCAGCGTACCCACAGCGGCGCCCTCTCTCTCTGCCAACCTGT 240
 Qy 2581 GCGTTTTCATCTGTGTCTCAGACCAACCGCGCTCCAGTCCACAGAGCAGAGAGGG 2640
 Db 241 GCGTTTTCATCTGTGTCTCAGACCAACCGCGCTCCAGTCCACAGAGCAGAGAGGG 300
 Qy 2641 ACAATGCCCGCCACCATCAGAAATGCGTGTATCTGCGCAGACTTGAATCTTCAATGTGCCA 2700
 Db 301 GCGGTGCCCGCGCACCTGAAAGTGTCTGTCTCTTCCAGATTTGAGCCCTCCATGTGCCA 360
 Qy 2701 CAGGTGTGTGCGAGTTCTCTCCAGACTGTGGGAGCTGTGGGAGTGAAGGGGAGAGCTTG 2760
 Db 361 CCAGGTGTGTGCGAGTTCTCTCCAGACTGTGGGAGCTGTGGGAGTGAAGGGGAGAGGCTC 420
 Qy 2761 TGTGAACTGTGATCGGAACCTGCGCAGAGCTACTGTTTCCCTTGGAGCTGCTGGGGGCTGT 2820
 Db 421 TCGGGATGTGATCAGAACCACTGCGCAGAGCTGTATTCCTTGGAGCTGCTGGGGGCTGT 480
 Qy 2821 GGACAGGAACCTGGAGAGGGTGGAGGAGCTGCGAGCTGCGAGCTGCTGTTGTGGGGCTGTG 2880
 Db 481 GGACAGAGCTGGAGAGGGTGGAGGAGCTGCGAGCTGCGAGCTGCTGTTGTGGGGCTGTG 540
 Qy 2881 ATCGAGGAGAGGGTGGAGGAGGGTGGAGGAGCTGCGAGGAGGAGGAGGAGGAGGAGG 2940
 Db 541 ATCGAGGAGAGGGTGGAGGAGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 Qy 2941 TTTGCTGTAGTTTATGCGCCCTTTTGGCACTCACTACCCCAACCTGCTGGGCAATG 3000
 Db 601 TTTGCTGTAGCTCTGCGCTTGGCACTCACTATCCCAACCTGCTGGGCGGCACTG 660
 Qy 3001 AAGACTCAGTGTGAGAAACCTTCCGCTGTGCGCGCTGCTCCATAGCCCTCTGCTCAT 3060
 Db 661 AAGACACACAGTGTGAGAAACCTTCCGCTGTGCGCGCTGCTCCATAGCCCTCTGCTCAT 720
 Qy 3061 CTGGATTAACCTGAAACGCGCACCGAGCGCTCCATACAGGAGAGAGGCGCTTACAGTGC 3120
 Db 721 CTGGATTAACCTGAAACGCGCACCGAGCGCTCCATACAGGAGAGAGGCGCTTACAGTGC 780
 Qy 3121 CTCTGTGCGTATGCTGTGGCAACCTGCGCAACCTCAAGCGTCAATGCTGCGATCCACTCT 3180
 Db 781 CTCTGCGCTTATGCTGTGGCAATCTGCGCAACCTCAAGCGTCAATGCTGCGATCCACTCT 840
 Qy 3181 GGTGCAACAACTTTTGGGTGTAGCCCTTGGCACTACAGTGTGCAACAGAGTATGAACCTC 3240
 Db 841 GGTGCAACAACTTTTGGGTGTAGCCCTTGGCACTACAGTGTGCAACAGAGTATGAACCTC 900
 Qy 3241 AAACGCTCATATGCTCGGACACACAGGGGAGAGGCGCTTCCGCTGTGCGCACTGCGCCCTAT 3300
 Db 901 AAACGCTCATATGCTCGGACACACAGGGGAGAGGCGCTTCCGCTGTGCGCACTGCGCCCTAT 960

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

QY 3301 ACCACAGGCCACTGGGACAACTACAAGCGTCATCAGAAGGTGTCATGGCCATGTTGGAGCA 3360
 Db |||||
 QY 961 ACCACGGGCCACTGGGACAACTACAAGCGCCACACAGAAGGTGTCATGGCCACGGTGGGCA 1020
 Db |||||
 QY 3361 GGAGGGCTGGTCTCTCTGCGCCCTGAGGGCTGGGCCCACTCATAGCCACCCCTCTGTT 3420
 Db |||||
 QY 1021 GGAGGGCTGGTCTCTCTGCGCTCTGAGGGCTGGGCCCACTCATAGCCACCCCTCTGTT 1080
 Db |||||
 QY 3421 TTGAGCACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTCATTACAGAC 3480
 Db |||||
 QY 1081 TTGAGCTCTCGGGGCCACACAGCCCTGGGACTGCTGGCAGCGGGCTGTCCACACAGAC 1140
 Db |||||
 QY 3481 TCACCTTGAACCTAAGTCTGTTTACCTGGGCTCTAGGAATTAGCCCTATGCTCCT 3540
 Db |||||
 QY 1141 TCATCTGAACCTAGTCTCTCTC-----CCC 1167
 QY 3541 GCATTTTATACAAATGAACCTAGAAACCAACCTTTCCCTTTCTCCCGCTGGTCAGGGGCT 3600
 Db |||||
 QY 1168 ATGTTTATACAGACGGACACAGAGCCACTT-----TTTCTCCCGCTGGCCAGGGGCT 1223
 Db |||||
 QY 3601 CCACACAGACTAAGCTAGGCACTATATGGACAGCCCTGAATCCCATGGTCAGGGGG-CCA 3659
 Db |||||
 QY 1224 CCACACAGACTAAGCTAGGCACTATAGGACAGCCCAACCCCATGGCGGGGGGGCCCA 1283
 Db |||||
 QY 3660 TATAGCCAGGGGA-CTTGTCTTAGCTCACTACAGATGAGCTAAGTATTAGGGGCTT 3718
 Db |||||
 QY 1284 TATGACCCAGGGGACCTTGGCTTGAAGGCACTTCAAGAGCTCAGTAGAGGGGCTT 1343
 Db |||||
 QY 3719 GGATTACCGGCCTGCTCCAGAGGGCTATGGATGAACCTGGTTGGAG-CTGCCCCAGCCT 3777
 Db |||||
 QY 1344 GTATTCACTCCACTGCCCCCAGGGGCTGTGGACAAACCGGCTGGGGGACTGCCCCAGCCT 1403
 Db |||||
 QY 3778 TTTAC-----TGTTTAACTTTTTCAGTGTCTTTATATAAGGAACACTAAC 3826
 Db |||||
 QY 1404 CCCACCTGTTTATTAACTTTTTCAGTGTCTTTATATAAGGAACACTAAC 1456
 Db |||||

Search completed: May 14, 2004, 03:12:30
 Job time : 14677 secs

D.T.B: 4/28/98

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 15:37:55 ; Search time 251 Seconds
(without alignments)
8505.562 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 ggaaggttagactttattc.....gaaaaaaaaaaaaaaaaa 3847

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815.2	21.2	1558	4	US-09-620-312D-32
2	615.8	16.0	1824	2	US-08-822-701-1
3	615.8	16.0	1824	3	US-08-935-855-1
4	615.8	16.0	2081	3	US-08-935-855-21
5	270.2	7.0	1890	3	US-08-935-855-19
6	206.2	5.4	235	4	US-09-016-434-1049
7	113.8	3.0	2678	4	US-09-016-434-1229
8	111.8	2.9	2416	4	US-09-016-434-1058
9	110.4	2.9	524	4	US-08-833-381-427
10	100.4	2.6	3536	3	US-09-418-640-3
11	100.4	2.6	3720	1	US-08-074-967-1
12	100.4	2.6	3720	2	US-08-553-541B-1
13	100.4	2.6	3720	3	US-09-268-202-1
14	100.4	2.6	3720	5	PCT-US94-06669-1
15	96.4	2.5	2769	4	US-09-620-312D-309
16	94.8	2.5	2680	3	US-08-063-035-1
17	94.8	2.5	2992	4	US-09-362-123A-3
18	93	2.4	4468	4	US-09-620-312D-243
19	92	2.4	605	4	US-09-621-976-2809
20	92	2.4	1033	4	US-09-620-312D-500
21	90.2	2.3	639	4	US-09-833-381-1188
22	90	2.3	2666	4	US-09-881-578A-1
23	90	2.3	3186	4	US-09-016-434-1390
24	89.4	2.3	1892	2	US-08-933-750C-66
25	89.4	2.3	1892	3	US-09-234-613-66
26	87.6	2.3	644	4	US-09-833-381-819
27	85.8	2.2	128779	4	US-09-497-855A-38

c 27

ALIGNMENTS

RESULT 1
US-09-620-312D-32
; Sequence 32, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 32
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(1174)
US-09-620-312D-32

Query Match 21.2%; Score 815.2; DB 4; Length 1558;
Best Local Similarity 85.7%; Pred. No. 1.4e-201;
Matches 1004; Conservative 0; Mismatches 123; Indels 45; Gaps 7;

QY 2674 CTGCCAGATTGAGTTTCATGTGCCACGAGTGTGCCAGTTTCTGCGCAGACTGTGGG 2733
DB 425 CCACAGATTGAGCTTCATGTGCCACGAGTGTGCCAGTTTCTGCGCAGACTGTGGG 484

Sequence 817, App
Sequence 1337, App
Sequence 689, App
Sequence 1399, App
Sequence 809, App
Sequence 163, App
Sequence 130, App
Sequence 586, App
Sequence 1406, App
Sequence 1401, App
Sequence 431, App
Sequence 172, App
Sequence 44, Appl
Sequence 1177, App
Sequence 911, App
Sequence 43, Appl
Sequence 821, App
Sequence 1054, App

QY 2734 CAGTCGCGGCTGAAGGGGAGAGCTTGTGTGAACTCGATCCGAACCACTGCCAGAGCTA 2793
DB 485 CAGTCGCGGCTGAAGGGGAGAGGCTCTGCGGGAGCTGGATCAGAACTGCCAGAGAGCTG 544
QY 2794 CTGTTCCTTCGACCTGCGCGGCTGTGGAACAGAACTGGAGGAGGCTGAGGGCAGCAGG 2853
DB 545 CTATTCCTTCGACCTGCGCGGCTGTGGAACAGAGCTGGAGGAGGCTGAGGGTAGTCGG 604
QY 2854 CTGGAGAGCTGCCATGTGTGGGCTGTGCATGCGAGGAGGCTGAGGGGTGCCACTGGG 2913
DB 605 CTGGAGAGCTGCCATGTGTGGGCTGTGCATGCGAGGAGGCTGAGGGGTGCCAGTGGG 664
QY 2914 GSAACCCAGGGCCCTGTGACAAAGCTTGTGCTGTAGTTATGCCCCCTTTGCCACTCAC 2973
DB 665 GGGCCCCAGGGCCCTGACAAAGCTTGTGCTGTAGCTCTGCCCTTGTGCCACTCAC 724
QY 2974 TACCCCAACCACTGGCTGGGCAATGAGAGCTCTACAGTGTGAGAACTTCCGCTGT 3033
DB 725 TATCCCAACCACTGGCCCGGCACATGAGACACACAGTGTGAGAACTTCCGCTGC 784
QY 3034 GCCCGCTGTCCATAGCCTCTGCTCATCTGATTAACCTGAAACGCACACGCGGTCCAC 3093
DB 785 GCCCGCTGTCTTATGCTCTGCTCATCTGATTAACCTGAAACGCACACGCGGTCCAT 844
QY 3094 ACAGAGAAAGCCCTACAAGTGGCCCTCTGTTCGATGCTGTGGCAACTGGCCAAAC 3153
DB 845 ACAGAGAAAGCCCTACAAGTGGCCCTCTGTTCGATGCTGTGGCAACTGGCCAAAC 904
QY 3154 CTCAGAGCTCATGCTGCGATCCTCTGCTGTGAGCAAACTTTCGCTGTAGCCTTTGCAAC 3213
DB 905 CTCAGAGCTCATGCTGCGATCCTCTGCTGTGAGCAAACTTTCGCTGTAGCCTTTGCAAC 964
QY 3214 TACAGCTGCAACAGAGTATGAACCTCAACCGTCTATGCTGCGACACACGGGCGAGAG 3273
DB 965 TACAGCTGCAACAGAGTATGAACCTCAACCGTCTATGCTGCGACACACGGGCGAGAG 1024
QY 3274 CCTTCGCTGTGCACTGCGCTCTATACACAGCCACTGGGCAACTACAGGCTCAT 3333
DB 1025 -CCTTCGCTGTGCACTGCGCTCTATACACAGCCACTGGGCAACTACAGGCTCAT 1083
QY 3334 CAGAGGTGTCATGGGCTGCTGCTGAGCAGAGGCTGCTCTCTGCCCCCTGAGGGCTGG 3393
DB 1084 CAGAGGTGTCATGGGCTGCTGCTGAGCAGAGGCTGCTCTCTGCCCCCTGAGGGCTGG 1143
QY 3394 GCGCCACTCATAGCCACCTCTGCTTTGAGCACTCGGGGTCAGAGCCCTGGGCTCT 3453
DB 1144 GCGCCACTCATAGCCACCTCTGCTTTGAGCACTCGGGGTCAGAGCCCTGGGCTCT 1203
QY 3454 ACTGTGAGCAGGCTCTTCTCATTCAGACTCACCTTGAACTAACTAGGTTCTTTTACCTGG 3513
DB 1204 GTTGGAGCGCGGCTGTCCACAGACTCTCTGAC----- 1241
QY 3514 GCTCTAGGAATTAGCCCTATGCTCTGCAATTTTACAAATGAACTAGAAACCACTTT 3573
DB 1242 -----TAGTCTCTTCTTCCCATGTTTATACAGAGGACCAAGGACCACTT- 1289
QY 3574 CCTTCTCCCGCTGCTGAGGGCTCCACAGACTAACTAGGCTATATGAGCA 3633
DB 1290 ----TCTCTCCCGCTGCGCAGGGCTCCACAGACTAACTAGGCTATATGAGCA 1346
QY 3634 GCTGTAATCCCTATGCTGAGGGG-CCATATAGACCAAGGGA-CTTGTCTTAGCTCACGTA 3691
DB 1347 GCCCAACCCATAGGGGGGGGGCCCATATGGAACAGAGGACCTTGGCTTGACTGAGCA 1406
QY 3692 CCAGATGAGCTAGTATGAGGCTTGTGATTCAGCCACTGCTCCCAAGGCTATGGA 3751
DB 1407 CTTACGAGCTAGTGAAGGGCCCTGTATTACCTCCTGCTGCCCCCHAGGGCTGTGA 1466
QY 3752 TGAACCTGGTTGGAG-CTGCCAGCCCTTTTAC-----TGTTTTAACTTATTTCAGTCTTT 3806
DB 1467 CAAACCGCTGCGGGAGCTGCCAGCCCTCCACCTGTTTATTAACTTATTTCAGTCTTT 1526
QY 3807 ATATTAAGAAACACTTAACAGAAAAAAA 3838

DB 1527 ATATTAAGAAACACTTAACAGAAAAAAA 1558

RESULT 2

US-08-822-701-1
Sequence 1, Application US/08822701

Patent No. 5976853

GENERAL INFORMATION:

APPLICANT: Guthridge, Mark

APPLICANT: Basilio, Claudio

TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE

TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

STREET: Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/822,701

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1049-1-002 N

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1824 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-822-701-1

Query Match

16.0%; Score 615.8; DB 2; Length 1824;

Best Local Similarity 99.7%; Pred. No. 8.8e-150;

Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 373 AGTGTGTAATCAGTCTGCGCCAGACACTTCTGGGATGTACAGGTTGACAACA 432

DB 1187 AGTGTGTAATCAGTCTGCGCCAGACACTTCTGGGATGTACAGGTTGACAACA 1246

QY 433 TGAGTGCATCATCTTTGTTCAAGCCCGAAACACAGTAGAGCTTCAGGCAGAGTG 492

DB 1247 TGAGTGCATCATCTTTGTTCAAGCCCGAAACACAGTAGAGCTTCAGGCAGAGTG 1306

QY 493 GCAAGAGAACTGGAGGAGGCTGTCCACGAGGGGCTGAAGACACCGGCAACAGTG 552

DB 1307 GCAAGAGAACTGGAGGAGGCTGTCCACGAGGGGCTGAAGACACCGGCAACAGTG 1366

QY 553 ACAAAAGAGAGGCTCAGAGGAGTATAGTGTCAACCGGACCTGCCATGTGGACTGTTTT 612

DB 1367 ACAAAAGAGAGGCTCAGAGGAGTATAGTGTCAACCGGACCTGCCATGTGGACTGTTTT 1426

QY 613 CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCCTAGCAGTGGGTAT 672

DB 1427 CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCCTAGCAGTGGGTAT 1486

Qy	2405	GCCGCTGCCATACGNTCAGCCCGAGTTGCTCAACTGACGCGACATACCGGCAACCATA	2464
Db	4	GC CGCTGCCCTACGNTCAGCCCGAGTTGCTCAACTGACACGACATACCGNACCCACA	63
Qy	2465	CTGGCGGAGAGCCCTACGGTTGTCCCCACTGCGCCCTTTGCTGCGAGAGCCTGGGCAACC	2524
Db	64	CTGGCGGAGAGCCCTACCGCTGTCCCACTGCCCTTTGCTGCGAGAGCCTGGGCAACC	123
Qy	2525	TGAGCGGCGATCAGCGCACCCACACAGGGCCCTCCCACTCCCTCCCTGCCCAACCTGTGGCT	2584
Db	124	TGAGCGGCGATCAGCGTACCCACGAGGGCCCCCACTCTCTCCTGCCCGACCTGTGGCT	183
Qy	2585	TTCCATGTGTGTCTCCAGCAACACCGGCTCCCGAGTCCACAGAGCAGGA	2636
Db	184	TCCGTGTGTGTACTCCAGACAGCCCGGCGCTTCCAGTCCCAACAGAGCAGGA	235

	Query Match	3.0%;	Score 113.8;	DB 4;	Length 2678;
	Best Local Similarity	56.6%;	Pred. No. 2.1e-19;		
	Matches 211;	Conservative 0;	Mismatches 162;	Indels 0;	Gaps 0;
QY	2925	CCGCGGTGACAAAGGCTTCGCTGTAGTTATGCCCCCTTTGGCACTCACTACCCCAACCA	2984		
DB	1864	CACGGGCGACAGTCCTTTGGCTGGTCGAGTGGGGGAGCGCTTCGGCGCGGCTCAGT	1923		
QY	2985	CTTGGCTCGGCACATGAAGACTACAGTGTGTGAGAAACCTTCGGCTGTGCCGCTGTCC	3044		
DB	1924	GCTGCTCGACGACCGGCGCGGTGCACATGTGGCAGCGGCCCTTCGCCCTGTGCCGAGTGGG	1983		
QY	3045	ATACGCTCTGCTCATCTGGATACCTGAAACGGCACACAGCGGCTCCACACAGGAGAAA	3104		
DB	1984	CCAGAGCTTCGGGACGCGCTCAACCTTGACGACACACGGGCGCATCCACACGGGGAGCG	2043		
QY	3105	GCCCTACAGTGCCCGCTCTGTCGGTATGCGTGTGGGCAACTCGGCCAACCTCAAGCGTCA	3164		
DB	2044	GCCCTTCGCTTCGGCGGAGTGTGCAGAGCCCTTCGCGACGGGCTACGCTCAGCGACCA	2103		
QY	3165	TGTCGGCATCCACTCTGGTGACAAACCTTTTCGGTGTAGCGCTTTGGAACCTACAGCTGCAA	3224		
DB	2104	TCTCCGGGTACACACGGGCGAGAAACCTTTGCCCTGTGCCCGGAGTGTGGCCACGCTTCAG	2163		

QY 3225 CCAGAGTATGAACCTCAACGTCATATGCTGGACACACGCGGAGAGCCCTTCCGCTG 3284
 Db 2164 CCAGGCGCTCAAGCTACCGCTCATCAGAGACACACACCGCGGAGAGCCCTTACCACTG 2223
 QY 3285 TGCCACCTGGCC 3297
 Db 2224 CGGTGAGTGGGC 2236

RESULT 8
 US-09-016-434-1058
 ; Sequence 1058, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1058:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2416 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g1050529
 ; US-09-016-434-1058

QY 2930 GTGACAAAGCCTTGCTGTAGTTATGCGCTTTGCGCACTCACTACCCCAACCACTGG 2989
 Db 761 GCGAGAAGGCTTACAAGTGGGATGCGCAAGGCTTCACTGGAGCAACCACTTC 820
 QY 2990 CTGGGACATGAAGACTCAGAGTGGTGAAGAACCTTCCGCTGTGCCGCTGCCATAG 3049
 Db 821 TGGAGACACCGCGCATCCACCGCGGAGAGCCCTTCTCTCGGGGAGTGCGGAAG 880
 QY 3050 CCTCTGCTCATCTGGATAACCTGAAGCGGACACGCGCGTCCACACAGGAGAAAGCCCT 3109
 Db 881 CCTTCAGCTGCCACTCGTCCCTCAACCTGACACGCGCATCCACACGCGGCGCGCCCT 940

Query Match 2.9%; Score 111.8; DB 4; Length 2416;
 Best Local Similarity 54.2%; Pred. No. 6.6e-19;
 Matches 227; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 2995 CACATGAAGACTCAGAGTGGTGAAGAACCTTCCGCTGTGCCGCTGTCCATAGCCTCT 3054
 Db 512 CACAGGAGGTCACACACAGGAGACAGACCCCTTTCAATGTGATCTCTGTGCGAAGCGCTC 453
 QY 3055 GCTCATCTGGATACCTGAAAGCGCACAGCGGCTCCACACAGGAGAGAAAGCCCTACAAG 3114
 Db 452 TTGAGCGCTCAGACCTCCGAGTTCACAGGAGTCCACACTGGGAGAGGCGCTTACATG 3393
 QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTGTATGGTTCGATC 3174
 Db 392 TGTGAGCTCTGCCAAAGCGGTTCCGCCACGAGTCCACCTTACAGGGCCACAGAGGATC 333
 QY 3175 CACTGTGGTGAACAAACCTTTTCGGGTAGCTTTGCGAATCTACAGCTGCAACCCAGAGTATG 3234
 Db 332 CACACCGGGGAGAGGCGGCTTTAAATGTAATATCTGAGCAAGATTTTCAGCCACAGGGG 273
 QY 3235 AACCTCAAAACCTCATATGCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCCTGC 3294
 Db 272 AACCTGAAGCTTCCAGCGGCAACCTTCCGAGAGAGGCGCTTACAAATGTCCACCTGT 213
 QY 3295 GCCTATACACAGCGCACTGGGACAACTACAGCGTCTATCAGAA 3338
 Db 212 CAAAAGGCGCTTCCGCTAGCTGGGAGACATTCAAGCGTCACTGAA 169

RESULT 10
 US-09-418-640-3
 ; Sequence 3, Application US/09418640
 ; Patent No. 6140125
 ; GENERAL INFORMATION:
 ; APPLICANT: Jennifer K. Taylor
 ; APPLICANT: Lex M. Cowser

QY 3110 ACAAGTSCCCCTCTGTCCGTATGCTGTGCAACCTGCGCAACCTCAAGCGTCAATGTC 3169
 Db 941 ACAAGTSCAGCGCTTCGAGAGAGGCTTTCAGCTGCACTGCTGCTGCTCAAGTCACTGC 1000
 QY 3170 GATCCACTCTGTGACAAACCTTTTCGGGTAGCTTTTGGAACTACAGCTGCNACCAGA 3229
 Db 1001 GGGTGACACACCGCGGAGAGCCCTTACCGGTGGCGAGTGGCGCAAGGCGCTTCAACAGC 1060
 QY 3230 GTATGAACCTCAAAACGTCATATGCTGCGACACACACGCGGCGAGAGCCCTTCCGCTGTGCCA 3289
 Db 1061 GTACACACTCAGACCGCCACCGCATCCACACGCGGCGAGAGCCCTTACCAGTGGGCT 1120
 QY 3290 CTTGGCGCTTATACACAGGCACTGGGACAACTACAGCGTATCAGAGTGTGATGCG 3348
 Db 1121 CTTGGCGCAGGCGCTTCACTGCACTCCCTGCGTATGCGTATGAGAGATCCACAGC 1179

RESULT 9
 US-09-833-381-427/C
 ; Sequence 427, Application US/09833381
 ; Patent No. 6672186
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; PRIOR FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 427
 ; LENGTH: 524
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-833-381-427

Query Match 2.9%; Score 110.4; DB 4; Length 524;
 Best Local Similarity 57.5%; Pred. No. 6.8e-19;
 Matches 198; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 2995 CACATGAAGACTCAGAGTGGTGAAGAACCTTCCGCTGTGCCGCTGTCCATAGCCTCT 3054
 Db 512 CACAGGAGGTCACACACAGGAGACAGACCCCTTTCAATGTGATCTCTGTGCGAAGCGCTC 453
 QY 3055 GCTCATCTGGATACCTGAAAGCGCACAGCGGCTCCACACAGGAGAGAAAGCCCTACAAG 3114
 Db 452 TTGAGCGCTCAGACCTCCGAGTTCACAGGAGTCCACACTGGGAGAGGCGCTTACATG 3393
 QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTGTATGGTTCGATC 3174
 Db 392 TGTGAGCTCTGCCAAAGCGGTTCCGCCACGAGTCCACCTTACAGGGCCACAGAGGATC 333
 QY 3175 CACTGTGGTGAACAAACCTTTTCGGGTAGCTTTGCGAATCTACAGCTGCAACCCAGAGTATG 3234
 Db 332 CACACCGGGGAGAGGCGGCTTTAAATGTAATATCTGAGCAAGATTTTCAGCCACAGGGG 273
 QY 3235 AACCTCAAAACCTCATATGCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCCTGC 3294
 Db 272 AACCTGAAGCTTCCAGCGGCAACCTTCCGAGAGAGGCGCTTACAAATGTCCACCTGT 213
 QY 3295 GCCTATACACAGCGCACTGGGACAACTACAGCGTCTATCAGAA 3338
 Db 212 CAAAAGGCGCTTCCGCTAGCTGGGAGACATTCAAGCGTCACTGAA 169

RESULT 10
 US-09-418-640-3
 ; Sequence 3, Application US/09418640
 ; Patent No. 6140125
 ; GENERAL INFORMATION:
 ; APPLICANT: Jennifer K. Taylor
 ; APPLICANT: Lex M. Cowser

; TITLE OF INVENTION: ANTISENSE MODULATION OF BCL-6 EXPRESSION
 ; FILE REFERENCE: R15-0102
 ; CURRENT APPLICATION NUMBER: US/09/418,640
 ; CURRENT FILING DATE: 1999-10-15
 ; NUMBER OF SEQ ID NOS: 89
 ; SEQ ID NO 3
 ; LENGTH: 3536
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (328)..(2448)
 ; US-09-418-640-3

Query Match 2.6%; Score 100.4; DB 3; Length 3536;
 Best Local Similarity 54.3%; Pred. No. 7.4e-16;
 Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
 QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCC 2979
 Db 1942 CAGACCCACAGTGAACAAACCCCTACAAGTGTGACCGCTGCCAGGCTCTTCCCGTACAAG 2001
 QY 2980 AACCACTCGCTGCGACATGAAGACTCACAGTGTGAGAAACCCCTCCGCTGCGCCG 3039
 Db 2002 GGCACCTCGCGACCAAGACCGCTCCATCCGTTGAGAAACCCCTATCTGTTGCAATC 2061
 QY 3040 TGTCCATACGCTCTGTCTCATCTGATTAACCTGAACCGGACCGCTCCACACAGGA 3099
 Db 2062 TGTGGGGCCAGTTCAACCGGCGACCAACCTGAACCCACACACTCGAATTCATCTGGA 2121
 QY 3100 GAAGAGCCCTCAAGTGGCCCCCTCTGTCCGTATCGCTGTGCGACCTGGCCACCTCAAG 3159
 Db 2122 GAGAAGCCCTCAATGGAACCTGCGAGCCAGATTTGTACAGTGGGCCACCTCCGT 2181
 QY 3160 CTTATGTGCGATCCACTCTGTGACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGC 3219
 Db 2182 GCCCATGTGCTTATCCACACTGTGAGAAAGCCCTATCCCTGTGAATCTGTGGCACCCGT 2241
 QY 3220 TGCACACAGATGAACTCAAGCTCATATGCTGCGACACACGCGGAGAGGCCCTTC 3279
 Db 2242 TTCGGGACCTTCAGACTCTGAGAGCCACTCGAATCCACACAGAGAGAAACCTTAC 2301
 QY 3280 CGCTGTGCCACCTG 3293
 Db 2302 CATTGTGAGAAGTG 2315

RESULT 11
 US-08-074-967-1
 ; Sequence 1, Application US/08074967
 ; Patent No. 5641672
 ; GENERAL INFORMATION:
 ; APPLICANT: DALLA-FAVERA, RICCARDO
 ; APPLICANT: CHAGANTI, RAJU S.K.
 ; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 ; TITLE OF INVENTION: bcl-6
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: COOPER & DUNHAM
 ; STREET: 30 ROCKEFELLER PLAZA
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10112
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PATENT IN RELEASE #1.0, VERSION #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/074,967
 ; FILING DATE:
 ; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
 ; NAME: WHITE, JOHN P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 0575/43771
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 977-9550
 ; TELEFAX: (212) 664-0525
 ; TELEX: 422523 COOP UI
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3720 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 328..2445
 ; US-08-074-967-1
 Query Match 2.6%; Score 100.4; DB 1; Length 3720;
 Best Local Similarity 54.3%; Pred. No. 7.6e-16;
 Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
 QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCC 2979
 Db 1942 CAGACCCACAGTGAACAAACCCCTACAAGTGTGACCGCTGCCAGGCTCTTCCCGTACAAG 2001
 QY 2980 AACCACTCGCTGCGACATGAAGACTCACAGTGTGAGAAACCCCTCCGCTGCGCCG 3039
 Db 2002 GGCACCTCGCGACCAAGACCGCTCCATCCGTTGAGAAACCCCTATCTGTTGCAATC 2061
 QY 3040 TGTCCATACGCTCTGTCTCATCTGATTAACCTGAACCGGACCGCTCCACACAGGA 3099
 Db 2062 TGTGGGGCCAGTTCAACCGGCGACCAACCTGAACCCACACACTCGAATTCATCTGGA 2121
 QY 3100 GAAGAGCCCTCAAGTGGCCCCCTCTGTCCGTATCGCTGTGCGACCTGGCCACCTCAAG 3159
 Db 2122 GAGAAGCCCTCAATGGAACCTGCGAGCCAGATTTGTACAGTGGGCCACCTCCGT 2181
 QY 3160 CTTATGTGCGATCCACTCTGTGACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGC 3219
 Db 2182 GCCCATGTGCTTATCCACACTGTGAGAAAGCCCTATCCCTGTGAATCTGTGGCACCCGT 2241
 QY 3220 TGCACACAGATGAACTCAAGCTCATATGCTGCGACACACGCGGAGAGGCCCTTC 3279
 Db 2242 TTCGGGACCTTCAGACTCTGAGAGCCACTCGAATCCACACAGAGAGAAACCTTAC 2301
 QY 3280 CGCTGTGCCACCTG 3293
 Db 2302 CATTGTGAGAAGTG 2315

RESULT 12
 US-08-553-541B-1
 ; Sequence 1, Application US/08553541B
 ; Patent No. 582858
 ; GENERAL INFORMATION:
 ; APPLICANT: DALLA-FAVERA, RICCARDO
 ; APPLICANT: CHAGANTI, RAJU S.K.
 ; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 ; TITLE OF INVENTION: bcl-6
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: COOPER & DUNHAM LLP
 ; STREET: 1185 AVENUE OF THE AMERICAS
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: /08/553,541B
 FILING DATE: May 28, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX: 422523 COOP UI
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3720 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 328..2445
 US-08-553-541B-1

Query Match 2.6%; Score 100.4; DB 2; Length 3720;
 Best Local Similarity 54.3%; Pred. No. 7.6e-16;
 Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy	2920	CAGGGCCCTGTGCAAAAGGCTTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCC	2979
Db	1942	CAGACCACAGTGACAAACCTCAAGTGTGACCGCTGCCAGSCCTCTTCCGCTACAAG	2001
Qy	2980	AACCACTGCTCGGCATGAGACTCAGAGTGTGAGAACTTTCGCTGTGCCCGC	3039
Db	2002	GGCAACTCGCCAGCCACAGACCGTCAACCGGTGAGAAACCTTATGTTGCAACATC	2061
Qy	3040	TGTCCATACGCTCTGTCTGTGATTAACCTTGTGAGAACTTTCGCTGTGCCCGC	3099
Db	2062	TGTGGGCCCACTTCAACCGCCAGCCCACTGCAAGTGTGACCGCTGCCAGSCCTCTTCCGCTACAAG	2121
Qy	3100	GAAGAGCCCTACAGTGTGCAAAAGGCTTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCC	3159
Db	2122	GAGAGCCCTTCAAAATGCGAAACCTTGGAGAGCCAGATTTGTACAGTGTGCCCGCTTCC	2181
Qy	3160	CGTCATGTGCGCATCCACTCTGTGTGACAAACCTTTTGGTGTAGCCCTTTGCAACTACAGC	3219
Db	2182	GCCCATGTGCTTATCCACACTGTGTGAGAGCCCTATCCCTGTGAAATCTGTGCAACCGT	2241
Qy	3220	TGCAACAGAGTATGAACCTCAACCTCATATGCTGGGACACACGGGCGAGAGCCCTTC	3279
Db	2242	TTCCGGCACCTTCAGACTCTGAAGAGCCCTCGGAATCCACACAGGAGAAACCTTAC	2301
Qy	3280	CGCTGTGCCACCTG 3293	
Db	2302	CATTGTGAGAAGTG 2315	

RESULT 13
 US-09-268-202-1
 Sequence 1, Application US/09268202
 Patent No. 6174997
 GENERAL INFORMATION:
 APPLICANT: Dalla-Favera, Riccardo
 APPLICANT: Chaganti, Raju S.K.
 TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 TITLE OF INVENTION: bcl-6
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESS: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York

STATE: New York
 COUNTRY: United States of America
 ZIP: 10036
 COMPUTER READABLE FORM: disk
 MEDIUM TYPE: Floppy
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/268,202
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-2
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX: 422523 COOP UI
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3720 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 328..2445
 US-09-268-202-1

Query Match 2.6%; Score 100.4; DB 3; Length 3720;
 Best Local Similarity 54.3%; Pred. No. 7.6e-16;
 Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy	2920	CAGGGCCCTGTGCAAAAGGCTTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCC	2979
Db	1942	CAGACCACAGTGACAAACCTCAAGTGTGACCGCTGCCAGSCCTCTTCCGCTACAAG	2001
Qy	2980	AACCACTGCTCGGCATGAGACTCAGAGTGTGAGAAACCTTTCGCTGTGCCCGC	3039
Db	2002	GGCAACTCGCCAGCCACAGACCGTCAACCGGTGAGAAACCTTATGTTGCAACATC	2061
Qy	3040	TGTCCATACGCTCTGTCTGTGATTAACCTTGTGAGAACTTTCGCTGTGCCCGC	3099
Db	2062	TGTGGGCCCACTTCAACCGCCAGCCCACTGCAAGTGTGACCGCTGCCAGSCCTCTTCCGCTACAAG	2121
Qy	3100	GAAGAGCCCTACAGTGTGCAAAAGGCTTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCC	3159
Db	2122	GAGAGCCCTTCAAAATGCGAAACCTTGGAGAGCCAGATTTGTACAGTGTGCCCGCTTCC	2181
Qy	3160	CGTCATGTGCGCATCCACTCTGTGTGACAAACCTTTTGGTGTAGCCCTTTGCAACTACAGC	3219
Db	2182	GCCCATGTGCTTATCCACACTGTGTGAGAGCCCTATCCCTGTGAAATCTGTGCAACCGT	2241
Qy	3220	TGCAACAGAGTATGAACCTCAACCTCATATGCTGGGACACACGGGCGAGAGCCCTTC	3279
Db	2242	TTCCGGCACCTTCAGACTCTGAAGAGCCCTCGGAATCCACACAGGAGAAACCTTAC	2301
Qy	3280	CGCTGTGCCACCTG 3293	
Db	2302	CATTGTGAGAAGTG 2315	

RESULT 14
 PCT-US94-06669-1
 Sequence 1, Application PC/TUS9406669
 GENERAL INFORMATION:
 APPLICANT: Dalla-Favera, Riccardo
 APPLICANT: Chaganti, R.S.K.
 TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 TITLE OF INVENTION: bcl-6

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06669
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/074,967
FILING DATE: 09-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 43771-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 684-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
PCT-US94-06669-1

Query Match 2.6%; Score 100.4; DB 5; Length 3720;
Best Local Similarity 54.3%; Pred. No. 7.6e-16;
Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTGGCACTCACTACCCC 2979
DB 1942 CAGACCCACAGTGACAAACCCTACAAGTGTGACCGCTGCAGGCGCTCTTCCGCTACAAG 2001
QY 2980 AACCACTGGTCCGACATGAAGACTCAGAGTGTGAGAAACCTTCCGCTGTGCCCGC 3039
DB 2002 GGCACCTCGCCAGCCACAGACCGTCCATCCGGTGAGAAACCTTATCGTTGCAACATC 2061
QY 3040 TGTCCATACGCTCTGTCTCATCTGGATAACCTGAAACGCGACAGCGCGTCCACAGGA 3099
DB 2062 TGTGGGGCCAGTTTCAACCGGCGAGCCAACTGAAACCCACACTCGAATTCACCTGGA 2121
QY 3100 GAAAGCCCTACAGTGGCCCTCTGTCCGATGCTGCGTGGCACTGGCCACCTCAAG 3159
DB 2122 GAGAAAGCCCTACAAATGGAACCTGGGAGCCAGATTTGACAGTGGGCCACCTCCGT 2181
QY 3160 CGTACGTGCGCATCCACTCTGCGTGAACAACTTTTCGGTGTAGCCCTTTGCAACTACAGC 3219
DB 2182 GCCCATGTGCTTATCCACACTGGTGAGAGCCCTATCCCTGTGAAATCTGTGGCACCCGT 2241
QY 3220 TGCACACAGATATGAACCTCAACGCTATATGCTGCGACACAGCGCGGAGAGCCCTTC 3279
DB 2242 TTCGGCACCTTCAGACTCTGAGAGCCCACTGCGAATCCACAGAGAGAACTTAC 2301
QY 3280 CGCTGTGCCACTG 3293
DB 2302 CATTGTGAGAGTG 2315

RESULT 15
US-09-620-312D-309
; Sequence 309, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: dt_FL_genes Version 1.0
; SEQ ID NO 309
; LENGTH: 2789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (248)..(2659)
US-09-620-312D-309

Query Match 2.5%; Score 96.4; DB 4; Length 2769;
Best Local Similarity 55.8%; Pred. No. 7.1e-15;
Matches 184; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 2984 ACCTGGCTCGGCACATGAAGACTCAGAGTGGTGAGAAACCCCTTCCGCTGTGCCCGCTGTC 3043
DB 1626 ACCTGAAGGCCCACTGAAGATCCACATCGCTGACGGGCCCTCAAGTCCGAGAGTGTG 1685
QY 3044 CATAGGCTCTGTCTCATCTCGATTAACCTGAAACGGACACAGCGCTCCACACAGGAGAA 3103
DB 1686 GGAAGCAGTTTCCACCCTCAGGGAACCTGAAGCGGACCTTCGGATCCACAGCGGGAGA 1745
QY 3104 AGCCCTACAGTGCCCTCTGTCCGATGCTGCGCAACCTGGCAACCTCAAGCGTC 3163
DB 1746 AGCCCTACGTGTGATCCACTGCGACGACAGTTTGACAGCCCGGGCTCTGCAGCGGC 1805
QY 3164 ATGTGCGATCCATCTCGGTGAACAACTTTTCGGTGTAGCCTTTTGCACACTACAGTGCA 3223
DB 1806 ACGTCCGCAATTCACACAGGTGAGAGCCCATGCCAGTGTGTGATGTGCGGTAAAGCCCTCA 1865
QY 3224 ACCAGAGTATGAACCTCAAACTCATATGCTGGACACACAGCGCGGAGAGCCCTTCGGCT 3283
DB 1866 CCCAGGCCAGCTCCCTCATCGCCCACTGAGCCACACACCGGGGAGAGCCCTTACGTCT 1925
QY 3284 GTGCCACTGCGGCTTATACACAGGCCACT 3313
DB 1926 GCGAGCGCTGCGGCAAGAGATTGTCCTCAGT 1955

Search completed: May 14, 2004, 05:50:07
Job time : 266 secs

us-09-673-994a-7.rni

Fri May 14 09:10:45 2004

Result	Query	Match	Length	DB	ID	Description
No.	Score					

1885	GAGGAGAGATTCCAAAGACACCTCCCTCTTTAGTGTGATCTCAAGATTCCTTCGACGAA	1844
QY		
102	GAGGAGGAATCCCAAGACCCCTCCCTCTTCACTGTGATCTCAAGACTCCCTCGACGAA	161
Db		
1945	GTCTCTGGGGCCCTGGTGTGTGGAGAGCGATTGTCTACTAGCCCAAGATCTGGAGTTTGA	2004
QY		
162	GGACCCGGGGCCCTGGTATTGGAGAGTAGTATTGTCTACTAGCCCAAGATCTGGAGTTTGA	221
Db		
2005	GAGGAAGAGGAAGAGGATCAAGTGTACGCGCCCAACACCAAGCATCTGGGCTTTTGAGAG	2064
QY		
222	GAGGAGAGGAAGAGGAGGAAGCGCGCGCAACAGTCAACCACTCTAGGGCTTCGAGAG	281
Db		
2065	GACTTGAAGAGACTCTCAGGGGGCGCAGACTGTGAATTTCCTATGGGCTGAGTGACGAC	2124
QY		
282	GACTCGAGAGGAGACTCTCTGGGGCGCAGGCTGGGGCTTCCTATGGGCTGAGCAGCAT	341
Db		

ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: misc.feature					
OTHER INFORMATION: Incyte ID No. US20040033942A1 870100CB1					
US-10-221-625-121					
Query Match		38.9%;	Score 1498;	DB 13;	Length 1975;
Best Local Similarity		88.6%;	Pred. No. 0;		
Matches 1707; Conservative		0;	Mismatches 176;	Indels 44;	Gaps 6;
QY	1917	AGTGGATACTGAAGATTCCTTCGACGAAGGTCTCTGGGCCCTTGFGTTGGAGAGCGATT	1976		
DB	86	AGTGGATACTGAAGACTCCTCTCGACGAAGCACCCGGGCCCTTGTATTTGGAGAGTGATT	145		
QY	1977	GCTACTAGGCACAGATCTGGAGTTTTCAGAGSAGNAGGAGAGGATGAGGTGACGGCCA	2036		
DB	146	GCTACTAGGCCAGGATCTGGAGTTTTCAGAGSAGNAGGAGAGGAGGAGGAAAGCGCAGCGCA	205		
QY	2037	CACAGACCAGCTCATGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACC	2096		
DB	206	CAGTAGCACCTCATGGGCTTCGAGAGAGACTCGAAGAGAGACTCTCTGGGGGCCAGGCC	265		
QY	2097	TGCACTTCCTATGGGCTGAGTGACGACAGTCTGGGGCGGCGGCACACTTAAGTGCAGA	2156		
DB	266	TGGGCTTCCATATGGGCTGAGCAGCATGAGTCTGGGGCGGCGGCAGGCACTAAGTGCAGA	325		
QY	2157	GAGTGAAGTTGAGAACACAGCCAGAGGGTCCAGGGGAGGGCCAGGGGTGAGAGGCCAGGGCCC	2216		
DB	326	GAGTGAAGTTGAGAGGCCAGCCAGGGGTTCAGGGGAGGCTCAGGGGTGAGAGGCCAGGGCCC	385		
QY	2217	AGGCTGTCACTGTGTGGGGGGCCGACAGGTGAGGGGCCGTGTGTGGGGGACAGAGGGGG	2276		
DB	386	AGCCTGCCAGCTGTGTGGGGGGCCGACAGGTGAGGGGCCGTGTGTGGGGGACAGAGGGGGC	445		
QY	2277	GGGTGGGGGCCCCCGTGTGCCACAGGTTCATGTCATCATGCGCGCTGTGCGCTTTCGT	2336		
DB	446	GGGTGGGGGGCCCCGTGTGCCACAGGTTCATGTCATCATGCGCGCTGTGACCTTCGT	505		
QY	2337	GTCCCACTACTCGAGCCACTGGAAGCGGCAATGTCAGACACAACAGCGGGGAGAGCGGTT	2396		
DB	506	GTCCCACTACTCGAGCCACTGGAAGCGGCAATGTCAGACACAACAGCGGGGAGAGCGGTT	565		
QY	2397	CCGCTGTGGCGCTGCCCATACGCTCAGCGCCAGTTCGTCAAACCTCAGCGGACATACCCG	2456		
DB	566	CCGCTGTGGCGCTGCCCATACGCTCAGCGCCAGTTCGTCAAACCTCAGCGACATACCCG	625		
QY	2457	CACCCACTACTGGCGAGAAGCCCTACCGTTGTCCCACTGCCCTCTTTCCTTCGACAGCCT	2516		
DB	626	CACCCACTACTGGCGAGAAGCCCTACCGTTGTCCCACTGCCCTCTTTCCTTCGACAGCCT	685		
QY	2517	GGGCAACCTGAGGCGGCATACGCGCACCCACAGGGGCTCCCACTCTCTCTCTGCCAAC	2576		
DB	686	GGGCAACCTGAGGCGGCATACGCGCATACCCACGAGGGGCCCCCACTCTCTCTCTGCCAC	745		
QY	2577	CTGTGGCTTTCTGATGCTGCTCCACGACCAACCCCGGCTCCAGTCCCAACAGAGCAGGA	2636		
DB	746	CTGTGGCTTTCTGATGCTGCTCCACGACCAACCCCGGCTCCAGTCCCAACAGAGCAGGA	805		
QY	2637	GGGGACAATGCCCCGACGATCAGAAAATGGCGCTGATCTCTGCCAGATTGAGTCTTATGT	2696		
DB	806	GGGGCGGTGCCCCGGGACCTGAAGATGCTCTGCTCTCTTCCAGATTTGAGCCTCCATGT	865		
QY	2697	GCCACAGGTGGTGCCAGTTTCTGCCAGACTGTGGGCGAGCTGGGGGTGAAAGGGGAGAG	2756		
DB	866	GCCACAGGTGGTGCCAGTTTCTGCCAGACTGTGGGCGAGCTGGGGGTGAAAGGGGAGAG	925		
QY	2757	CTTGTGTGGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTTGGACCTGCCGGG	2816		
DB	926	CCTCTGGGGACTGGATCAGAAACCACTGCCAGAGCTGCTATTCCCTTTGGACCTGCCGGG	985		
QY	2817	CTGTGGACAGGAATCTGAGAGGGGTGAGGCGACAGGCTGGGAGCTGCCATGTGTGGCG	2876		
DB	986	CTGTGGACAGAGCTGAGAGGGGTGAGGGTAGTCTGGGTGGGAGCTGCCCTGTGGGCG	1045		

; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC06
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7625
 ; LENGTH: 9266
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-891-7625

Query Match 31.1%; Score 1195; DB 10; Length 9266;
 Best Local Similarity 81.3%; Pred. No. 0;
 Matches 1568; Conservative 0; Mismatches 216; Indels 144; Gaps 10;
 2003 AGAGGAGAGGAGAGGAGATGAAGGTGACGGCCACAAACACAGCTCATGGGCTTTGAGA 2062
 1441 AAGTGAAGAGGGAGAGGTTGGAGCCTGATTCCTTACCTATGCTGATGCTCTCTTTATGT 1500
 2063 GAGACTCTGAGGAGACTCTCAGGGGGCCAGACCTGGACTTCCCTATGGGCTGAGTGAGC 2122
 1501 CTATTTACACAGAGACTCTCTGGGGCCAGGCTTGGCTTCCCTATGGGCTGAGCGAG 1560
 2123 ACGAGTCTGGGGCGCGCGCCTAAGTGGGAGAGTGAAGTTGAGGAAACACGACGAGG 2182
 1561 ATGAGTCTGGGGCGCGCGCCTAAGTGGGAGAGTGAAGTTGAGGAGCCAGCCAGGG 1620
 2183 GTCCAGGGAGGCGGGGTGAGAGCCAGCCAGCCTGTGACGTGTGTGGGGGCGCGA 2242
 1621 GTCAGGGAGGCGGGGTGAGAGCCAGCCAGCCTGTGACGTGTGTGGGGGCGCGA 1680
 2243 CAGGTGAGGGCGCGGTGTGTGGGGCAGGAGGGGGGGTGGGGCGCGCGCTGCCCCCAG 2302
 1681 CAGGTGAGGGCGCGGTGTGTGGGGCAGGAGGGGGGGTGGGGCGCGCTGCCCCCAG 1739
 2303 GGTGTACTGTACTCATGCGGCTGTGCGCTTTCGTGTCCTACTACTGAGCCACTGAAGC 2362
 1740 GGTGTACTGTACTCATGCGGCTGTGCGCTTTCGTGTCCTACTACTGAGCCACTGAAGC 1799
 2363 GGCACATGCAGACACACAGCGGGGAGAGCGCTTCCGCTGTGCGCGTGGCCCATACGCT 2422
 1800 GGCACATGCAGACACACAGCGGGGAGAGCGCTTCCGCTGTGCGCGTGGCCCATACGCT 1859
 2423 CAGCCAGTTCGTCACCTGACGCGACATACCGCCACCTACTGCGGAGAGCGCTTACC 2482
 1860 CAGCCAGTTCGTCACCTGACGCGACATACCGCCACCTACTGCGGAGAGCGCTTACC 1919
 2483 GTTGTCCCACTGCGCCCTTTGCTGTCAGAGCCTGGGCAACCTGAGCGGCATCAGCGCA 2542
 1920 GCTGTCCCACTGCGCCCTTTGCTGTCAGAGCCTGGGCAACCTGAGCGGCATCAGCGTA 1979
 2543 CCCACAGAGGCGCTCCACTCTCCCTGCGCAACCTGTCGCTTTCGATGTGCTGCTCCAC 2602
 1980 CCCACAGAGGCGCGCCACTCTCCCTGCGCAACCTGTCGCTTTCGATGTGCTGCTCCAC 2039
 2603 GACCAACCGGCGCTCCACTGTCACAGAGAGGAGGAGCAATGCCCGC----- 2651
 2040 GACCAGCCGCGCTCCAGTCCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2099
 2652 ----- 2651
 2100 TAAGACACACAGGAGCAAGATCTTGGGACATGGGTGGCTGACCTTAGGAATGCTTGG 2159
 2652 -----ACGATCAGAAATGGCTGTGATCTCCAGAGCTTGTGAT 2688
 2160 ATTGGATTCAAGCCAGGCTTGTGTCACAGAGTCTGTCTCTCTCCAGATTTGAGC 2219
 2689 CTTTATGTGCCACAGGTGTGGTCCAGTCTTCTGCGCAGACTGTGGGAGCTGCGGGGTGAA 2748

Db 2220 CTCCATGTGCCACAGGT-GTCCAGTTTCTCTGCCAGACTGTGGGAGCTGCGGGTGAA 2278
 Qy 2749 GGGAGAGCTTGTGTGAACCTGATCCGAACCACTGCCAGAGCTACTGTTCCTTTGGACC 2808
 Db 2279 GGGAGAGGCGCTCTGCGGAGCTGATCAGAACCACTGCCAGAGCTGATTTCCCTTTGGACC 2338
 Qy 2809 TGGCGGGGCTGTGACAGGAACTGAGGAGGAGGTGAGGCGAGGAGCTGGGAGTGCATG 2868
 Db 2339 TGGCGGGGCTGTGACAGAGCTGAGGAGGAGGTGAGGAGTGCATG 2398
 Qy 2869 TGTGGGCGCTGACATGCGAGAGAGGCTGAGGAGGTTGCCACTGGGGGAGCCAGGAGCCCT 2928
 Db 2399 TGTGGGCGCTGACATGCGAGAGAGGCTGAGGAGGTTGCCAGTGGGGGGGCCACACCTGT 2458
 Qy 2929 GGTGACAAAGGCTTTGCTCTAGTTTATGCCCTTTGCCACTCACTACCCCAACCACTG 2988
 Db 2459 -----ACAAAGGCTTTGCTCTAGCTTGCCTTGCCTTTGGCACTCACTATCCCAACCACTG 2514
 Qy 2989 GCTCGGACATGAAGACTCACTAGTGTGAGAAACCTTCCGCTGTGCGCTGTCCATAC 3048
 Db 2515 GCGCGGACATGAAGACACACAGTGTGTGAGAGCCCTTCCGCTGCGCGCTGTCTCTAT 2574
 Qy 3049 GCCTCTGCTCATCTGGATTAACCTGAACGCGCACAGCGCTGCCACACAGGAGAAAGCC 3108
 Db 2575 GCCTCTGCTCATCTGGATTAACCTGAACGCGCACAGCGCTGCCACACAGGAGAAAGCC 2634
 Qy 3109 TACAAGTGCCTCTCTGCGTATGCTGTGGCAACCTGCGCAACCTCAAGCGTCACTG 3168
 Db 2635 TACAAGTGCCTCTCTGCGTATGCTGTGGCAATCTGCGCAACCTCAAGCGTCACTG 2694
 Qy 3169 CGCATCCACTCTGTGTGACAAACCTTTTGGTGTAGCCTTTGCACTACAGCTGCAAGCCAG 3228
 Db 2695 CGCATCCACTCTGTGTGACAAACCTTTTGGTGTAGCCTTTGCACTACAGCTGCAAGCCAG 2754
 Qy 3229 AGTATGAACCTCAAAAGCTCATCTGCGACACACGCGCGAGAGCCCTTCCGCTGTGCC 3288
 Db 2755 AGCATGAACCTCAAAAGCTCATCTGCGGACACACGCGCGAGAGCCCTTCCGCTGTGCC 2814
 Qy 3289 ACCTGCGGCTATACACAGGCGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 3348
 Db 2815 ACCTGCGGCTATACACAGGCGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2874
 Qy 3349 CATGTGTGAGAGAGGAGGCGCTGCTCTGCGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 3408
 Db 2875 CAGGTGTGAGAGAGGAGGCGCTGCTCTGCGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2934
 Qy 3409 CCACCTCTGTTTGTGAGCACTGCGGCTCCAGAGCCCTGCGGTGTCTGCTGAGGAGCTGAGG 3468
 Db 2935 CCACCTCTGTTTGTGAGCACTGCGGCTCCAGAGCCCTGCGGAGCTGCTGCGGAGCGGCT 2994
 Qy 3469 CTTGATTGAGACTCAGCTTGAACCTAGGTTCTTTTACCTGGGCTCTAGGAATAGC 3528
 Db 2995 GTCCACACAGACTCATCTGAAC-----TAGG 3021
 Qy 3529 CCTATGCTCTCTGCACTTTTATACAAATGAACCTAGAAACCACTTTCCTTCTTCCCGCCG 3589
 Db 3022 TCTTCTTCCCATGCTTTTATACAGCGAGCCAGAGCCAGCTT-----TTTCTCCCGCCG 3077
 Qy 3589 TGGTCAAGGCTTCCACACAGACTACCTAGGAGCTATATGAGCAGGAGCTGAGTCCCATGG 3648
 Db 3078 TGGCAGAGGCTTCCACACAGACTAAGCTAGGAGCTATATGAGCAGGAGCCAGCCCATGG 3137
 Qy 3649 TCAGGGGG-CCATATAGACCCAGGGGA-CTTGTCTTTAGCTCAGTACCGAGATGAGTAACT 3706
 Db 3138 CGGGGGGGCCATATGAGCCAGGAGCTTGTCTTGTGAGGAGCTTCAAGAGCTCACT 3197
 Qy 3707 GATTAGGCGCTTGTGATTCACCGCCCTGCTCCAGAGCTATGAGTACCTGTTGGAG 3766
 Db 3198 GAGAGGGCGCTGTATTCACCTCCACTGCGCGGAGCTGTGGAGAAACCGGCTGGGG 3257
 Qy 3767 -CTGCCAGCCCTTTTAC-----TGTTTAACTTATTTTCACTGCTTTTATATTAAGGAAACA 3821

Db 3258 ACTGCCAGGCTCCACCTGTTATTATTACTTATTTCAGTGTCTTTATTAATAAGGAACA 3317

QY 3822 CTAACAGA 3829

Db 3318 CTAACAAA 3325

RESULT 4

US-10-094-749-1126

; Sequence 1126, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOKYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094, 749

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1126

; LENGTH: 1456

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-094-749-1126

Query Match 29.0%; Score 1115.2; DB 16; Length 1456;

Best Local Similarity 87.4%; Pred. No. 0;

Matches 1305; Conservative 0; Mismatches 144; Indels 44; Gaps 6;

QY 2341 CACTACTCGAGCCACCTGAAGCGGCACATGCAACACACAGCGGAGAGAGCCGTTCCGC 2400

Db 1 CACTACTCGAGCCACCTGAAGCGGCACATGCAACACACAGCGGAGAGAGCCGTTCCGC 60

QY 2401 TGTGGCGGTGCCCATACGNCNTACGCCAGTTGCTTACCTGACGCGACATACCGGACCC 2460

Db 61 TGTGGCGGTGCCCATACGNCNTACGCCAGTTCGTCACCTGACACGACATACCGGACCC 120

QY 2461 CATACTGGCGAGAACCCCTACCGTTGTCGCCACTGCCCCCTTTGCTGCGAGCAGCGCTGGC 2520

Db 121 CACACTGGCGAGAACCCCTACCGTTGTCGCCACTGCCCCCTTTGCTGCGAGCAGCGCTGGC 180

QY 2521 AACCTGAGGGGATCAGCGGACCCACACAGAGGCTCCGACCTCCCTGCGGACACCTGT 2580

Db 181 AACCTGAGGGGATCAGCGGACCCACACAGAGGCTCCGACCTCCCTGCGGACACCTGT 240

QY 2581 GGCTTTTCGATGCTGCTCCAGGACCAACCCGCGCTCCGAGTCCCAACAGAGCAGAGGGG 2640

Db 241 GGCTTTTCGATGCTGCTCCAGGACCCACACAGAGGCTCCGAGTCCCAACAGAGCAGAGGGG 300

QY 2641 ACAATGCCCGGAGATCAGAAAATGGCTGATCCCTGCGAGATTCAGTCTTCATGTGCCA 2700

Db 301 GCGGTGCCCCGGCGACCTGAAGATGCTGCTCTTCCAGATTTTGAAGCTTCATGTGCCA 360

QY 2701 CCAGTGTGTGCCAGTTTCTGTCACAGCTGTGGGCACTGTGGGGTGAAGGGAGAGCTTG 2760

Db 361 CCAGTGTGTGCCAGTTTCTGTCACAGCTGTGGGCACTGTGGGGTGAAGGGAGAGCTTG 420

QY 2761 TGTGAACTGGATCCGAACCACTGTCAGAGCTACTGTTCCCTTGGAGCTGCCGGGCTGT 2820

Db 421 TGGGGACTGGATCAGAACCACTGTCAGAGCTGTATTCCCTTGGAGCTGCCGGGCTGT 480

QY 2821 GGACAGAACTGGAGAGGGTGAAGGCAAGGCTGGAGCTGGAGCTGCCATGTGTGGGCGTGC 2880

Db 481 GGACAGAGCTGGAGAGGGTGAAGGCTGGAGCTGGAGCTGGAGCTGCCATGTGTGGGCGTGC 540

QY 2881 ATGCCAGAGAGGCTGGAGGGGCTGCCCACTGGGGGACCCAGGGCCCTGTGTGAACAAGGC 2940

Db 541 ATGCCAGAGAGGCTGGAGGGGCTGCCCACTGGGGGACCCAGGGCCCTGTGTGAACAAGGC 600

QY 2941 TTTGCCCTGTAGTTTATGCCCCCTTTGCCACTCACTAACCCAAACCACTGGCTGGGACATG 3000

Db 601 TTTGCCCTGTAGCTCTGCCCCCTTTGCCACTCACTATCCCAACCACTGGCTGGGACATG 660

QY 3001 AAGACTCACAGTGTGAGAAACCCCTTCGCTGTGCGCTGTGCGCTGTGCTATACGCTCTGCTCAT 3060

Db 661 AAGACACACAGTGTGAGAAACCCCTTCGCTGTGCGCTGTGCGCTGTGCTATACGCTCTGCTCAT 720

QY 3061 CTGGATAACTGAAACCGGACCCAGCGCTCCACACAGAGAGAAAGCCCTACAAGTGGCCCC 3120

Db 721 CTGGATAACTGAAACCGGACCCAGCGCTCCACACAGAGAGAAAGCCCTACAAGTGGCCCC 780

QY 3121 CTCTGTCGATGCTGTGGGAACCTGGCCACCTCAAGCGCTCATGTGCGCATCCACTCT 3180

Db 781 CTCTGCCCCCTTATGCTGTGGCAACTGTCGCAACCTCAAGCGCTCATGTGCGCATCCACTCT 840

QY 3181 GGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACCAAGATGAACCTC 3240

Db 841 GGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACCAAGATGAACCTC 900

QY 3241 AAACGTCATATGCTGGGACACAGCGGCGAGAGCCCTTCGCTGTGCGCTGTGCGCTGTAT 3300

Db 901 AAACGTCATATGCTGGGACACAGCGGCGAGAGCCCTTCGCTGTGCGCTGTGCGCTGTAT 960

QY 3301 ACCACAGGCTGGGACAACTACAAGCTCATCAGAAAGTGCATGSCCATGTTGGAGCA 3360

Db 961 ACCACAGGCTGGGACAACTACAAGCTCATCAGAAAGTGCATGSCCATGTTGGAGCA 1020

QY 3361 GGAGGGCTGTCTCTGTCCTGAGGCTGGGCGCCACCTCATAGCCACCTCTGTT 3420

Db 1021 GGAGGGCTGTCTCTGTCCTGAGGCTGGGCGCCACCTCATAGCCACCTCTGTT 1080

QY 3421 TTGAGCACTCGGGGTCCAGCGCCCTGGGTGCTACTGTGTAGCAGGGCTCTTTCATTCAGAC 3480

Db 1081 TTGAGCTCTCGGGGCGCCACAGCCCTGGGACTGTGGCAGCGGGCTGTCCACACAGAC 1140

QY 3481 TCACCTTGAACTAACTAGGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCTCT 3540

Db 1141 TCACCTTGAACTAGGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCTCTCT 1167

QY 3541 GCATTTTATACAAATAAGAACTAGAAACCACTTTCCCTTTCTCCCGCTGTGTCAGGGCT 3600

Db 1168 ATGTTTTATACAGACGAGCCAGAACCTTT---TTTCTCCCGCTGTGTCAGGGCT 1223

QY 3601 CCACACAGCTAACCTAGGCTATATGAGCGCTGAATCCCATGTCAGGGG---CCA 3659

Db 1224 CCACACAGCTAACCTAGGCTATATGAGCGCTGAATCCCATGTCAGGGGCGGCCCA 1283

QY 3660 TATAGACCAAGGGGA---CTTGTCTTAGCTCACGTACCAAGTGAAGTGAAGTGAAGGCTT 3718

Db 1284 TATGACCAAGGGGA---CTTGTCTTAGCTCACGTACCAAGTGAAGTGAAGTGAAGGCTT 1343

QY 3719 GGAATTCACCGCACTGCTCCAGAGGCTATGATGAAGTGAAGTGAAGTGAAGGCTT 3777

Db 1344 GGAATTCACCGCACTGCTCCAGAGGCTATGATGAAGTGAAGTGAAGTGAAGGCTT 1403

QY 3778 TTTAC-----TGTATTACTATTTCAGTGTCTTTATATAAAGGAAACACTAAC 3826
 DB 1404 CCCACTGTTATTATTAACCTTATTCAGTGTCTTTATATAAAGGAAACACTAAC 1456
 RESULT 5
 US-10-037-270-32
 ; Sequence 32, Application US/10037270
 ; Publication No. US20030104529A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungling
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Tillinghast, John
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/10/037,270
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1104
 ; SOFTWARE: pc_fl_genes Version 1.0
 ; SEQ ID NO 32
 ; LENGTH: 1558
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (50)..(1174)
 US-10-037-270-32
 Query Match 21.2%; Score 815.2; DB 15; Length 1558;
 Best Local Similarity 85.7%; Pred. No. 1.6e-231;
 Matches 1004; Conservative 0; Mismatches 123; Indels 45; Gaps 7;
 QY 2674 CTGCCAGACTTGAGTCTTCATGTGCCACACAGGTGTGTCAGTTTCCTGCCAGACTGTGGG 2733
 DB 425 CCGACAGATTGAGCTTCATGTGCCACAGGTGTGTCAGTTTCCTGCCAGACTGTGGG 484
 QY 2734 CAGCTGCGGGTGAAGGGAGAGCTTGTGTGGAACTGGATCCGAAACCACTGCCAGAGCTA 2793
 DB 485 CAGCTGCGGGTGAAGGGAGAGGGCTCTGCGGGACTGGATCAGAAACCACTGCCAGAGCTG 544
 QY 2794 CTGTTCCCTTGACCTGCGGGGTGTGCAAGAACTGGAGGAGGTGAGGGCAGCAGG 2853
 DB 545 CTATTCCCTTGACCTGCGGGGTGTGCAAGAACTGGAGGAGGTGAGGGTAGTCCG 604
 QY 2854 CTGGAGCTGCCATGTGTGGGCGTGTGATGCGAAGAGGCTGAGGGGTGGCATCTGGG 2913
 DB 605 CTGGAGCTGCCATGTGTGGGCGTGTGATGCGAAGAGGCTGAGGGGTGGCATCTGGG 664
 QY 2914 GGACCCCGAGGCGCTGTGCAAAAGGCTTTCCTGTAGTTATGCCCTTTGCCACTCAC 2973
 DB 665 GGGCCCGAGGCGCTGTGCAAAAGGCTTTCCTGTAGTTATGCCCTTTGCCACTCAC 724
 QY 2974 TACCCCAACCACTGGCTCGGCACATGAACTACAGTGTGTGAGAAACCCCTCCGCTGT 3033

DB 725 TATCCCAACCACTGCGCCCGGCACATGAAGACACACAGTGTGTGAGAAGCCCTTCGCTGC 784
 QY 3034 GCGCGTGTCCATACCGCTCTGCTCATCTGGATACCTGAAACGGCACACAGCGCTCCAC 3093
 DB 785 GCGCGTGTCTTATGCTCTGCTCATCTGGATAAACCCTGAAACGGCACACAGCGCTCCAT 844
 QY 3094 ACAGGAGAAAAGCCCTACAAAGTGCCCTCTGTTCGTATGCTGTGGCAACCTGCGCAAC 3153
 DB 845 ACAGGAGAGAAGCCCTACAAAGTGCCCTCTGTTCGTATGCTGTGGCAACCTGCGCAAC 904
 QY 3154 CTCAAGGCTCATGCTGCGCATCCACTCTGGTGACAAACCTTTTCGGTGTAGCTTTGCAAC 3213
 DB 905 CTCAAGGCTCATGCTGCGCATCCACTCTGGTGACAAACCTTTTCGGTGTAGCTTTGCAAC 964
 QY 3214 TACAGTGCAAACAGAGTATGAACCTCAAAACGTCATATGCTCGCAACACAGCGCGCAGAG 3273
 DB 965 TACAGTGCAAACAGAGTATGAACCTCAAAACGTCATATGCTCGCGCACACAGCGCGCAGAG 1024
 QY 3274 CCCTTCCGCTGTGCCACCTGCGCTATACACAGGGCCTGCGGCACTTACAAAGCGTCAAT 3333
 DB 1025 -CCTTCCGCTGTGCCACCTGCGCTATACACAGGGCCTGCGGCACTTACAAAGCGGCAAC 1083
 QY 3334 CAGAAAGTGCATGCGCATGCTGAGCAGAGGGCTGTGCTCTCTGCGCTCTGAGGGCTGG 3393
 DB 1084 CAGAAAGTGCATGCGCATGCTGAGCAGAGGGCTGTGCTCTCTGCGCTCTGAGGGCTGG 1143
 QY 3394 GCGCCACCTCATAGCCCACTCTGTTTGGAGCACTCGGGGTCCAGACGCCCTGCGTGTCT 3453
 DB 1144 GCGCCACCTCATAGCCCACTCTGTTTGGAGCACTCTGCGGGTCCAGACGCCCTGCGGACT 1203
 QY 3454 ACTGTTAGCAGGGCTCTTCACTCAGACTCACCTTGAACCTAACTAGTGTCTTTTACCTGGG 3513
 DB 1204 GCTGGCAGCGGGCTGTGCCACAGACTCATCTGAAC----- 1241
 QY 3514 GCTCTAGGAATTAGCCCTATGCTCTGCACTTTTATACAAATGAACCTAGAAACCACTTT 3573
 DB 1242 -----TAGTCTCTTCTCCCATGTTTATACAGCGGACCTTACAGCGGACCTT- 1289
 QY 3574 CCCTTTCTCCCGCTGTGTCAGGGGTCCACACAGACTAACCTAGGACATATATGACCA 3633
 DB 1290 ---TTTCTCCCGCTGTGTCAGGGGTCCACACAGACTAACCTAGGACATATATGACCA 1346
 QY 3634 GCCTGAATCCCATGTCAGGGG-CCATATAGACACAGGGA-CTTGTCTTAGCTCACCTA 3691
 DB 1347 GCCCAACCCCATGTCAGGGGSCCCCATATGACACAGGACCTTGCCTTACTGAGCA 1406
 QY 3692 CCAGATGAGCTAAGTGAATTAGGGCTTGGATTACCGCCACTGCTCCAGAGGCTATGGA 3751
 DB 1407 CTTACAGAGCTCAGTGAGAAGGGCTCTGTATTACCTCCACTGCCCCCAGGGGCTGTGGA 1466
 QY 3752 TGAACCTGTTGGAG-CTGCGCCAGCTTTTAC-----TGTTTAACTTATTTCAGTGTCTT 3806
 DB 1467 CAACCCGGCTGGGGACTGCCAGCTCCACCTCCACCTGTTTATTAACTTATTTCAGTGTCTT 1526
 QY 3807 ATAATAAGGAACACATAACAGAAAAA 3838
 DB 1527 ATAATAAGGAACACATAACAAAAA 1558

RESULT 6

US-10-117-722-32
 ; Sequence 32, Application US/10117722
 ; Publication No. US20030219744A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 784CIP2BCIP
 ; CURRENT APPLICATION NUMBER: US/10/117,722

```

; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 32
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(1174)
; US-10-117-722-32

Query Match      21.2%; Score 815.2; DB 16; Length 1558;
Best Local Similarity 85.7%; Pred. No. 1.6e-231;
Matches 1004; Conservative 0; Mismatches 123; Indels 45; Gaps 7;

QY 2674 CTGCCAGACTTGAAGTCTTCTGATGCTCCACACAGGTGGTCCAGTTTCTCTGCCAGACTGTGGG 2733
DB 425 CCGACAGATTGACCTCCATGTGCCACCAAGTGTGCCAGTTTCTCTGCCAGACTGTGGG 484
QY 2734 CAGCTGGGGTGAAGGGAGAGCTTGTGTGGAACTGGATCCGAACCACTGCCAGACTA 2793
DB 485 CAGCTGGGGTGAAGGGAGAGGCTCTGCGGGGACTGGATCAGAAACCACTGCCAGACTG 544
QY 2794 CTGTTCCCTTGGACTGCTCGGGGCTGTGGACAGAACTGGAGGAGGCTGAGGGCAGCAGG 2853
DB 545 CTATTCCTTGGACTGCTCGGGGCTGTGGACAGAACTGGAGGAGGCTGAGGGTAGTCGG 604
QY 2854 CTGGAGCTGCCATGTGTGGCGCTGTGCATGCGAGAGGCTGAGGGGTTGCCACTGGG 2913
DB 605 CTGGAGCTGCCATGTGTGGCGCTGTGCATGCGAGAGGCTGAGGGGTTGCCACTGGG 664
QY 2914 GGACCCAGGGCCCTGGTGACAGAGGCTTTCCTGTAGTTTATGCCCTTTGCCACTCAC 2973
DB 665 GGGCCCAAGGGCCCGAGTGACAGAGGCTTTCCTGTAGTTTATGCCCTTTGCCACTCAC 724
QY 2974 TACCCCAACCACTGGCTCGGCACATGAAGACTCACAGTGTGTGAGAAACCTTCCGCTGT 3033
DB 725 TATCCCAACCACTGGCCCGGCACATGAAGACTGTGTGAGAAACCTTCCGCTGTG 784
QY 3034 GCGCGCTGTCCATACGCTCTGTCTCATCTGATACCTGAAACGCGCAGCGGCTCCAC 3093
DB 785 GCGCGCTGTCTCTGTCTCTCATCTGATACCTGAAACGCGCAGCGGCTCCAT 844
QY 3094 ACAGAGAGAAAGCCCTCAAGTGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAAAC 3153
DB 845 ACAGAGAGAAAGCCCTCAAGTGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAAAC 904
QY 3154 CTCAGAGCTCATGTGTGGCATCCACTCTGCTGTGACAAACCTTTCGCTGTAGCTTTGCCAC 3213
DB 905 CTCAGAGCTCATGTGTGGCATCCACTCTGCTGTGACAAACCTTTCGCTGTAGCTTTGCCAC 964
QY 3214 TACAGCTCAACACAGATGATGAACCTCAAGCTCATATGCTGCGACACACCGGCGAGAG 3273
DB 965 TACAGCTCAACACAGATGATGAACCTCAAGCTCATATGCTGCGACACACCGGCGAGAG 1024
QY 3274 CCGTTCCGCTGTGCCACCTTCCGCTTATACACAGGCCAATGAGCAATCAAGCTCAT 3333
DB 1025 CCGTTCCGCTGTGCCACCTTCCGCTTATACACAGGCCAATGAGCAATCAAGCTCAT 1083
QY 3334 CAGAGGTGCATGGCCATGGTGGAGCGAGGGGCTGCTCTCTGCGCCCTGAGGGCTGG 3393
DB 1084 CAGAGGTGCATGGCCATGGTGGAGCGAGGGGCTGCTCTCTGCGCCCTGAGGGCTGG 1143
QY 3394 GCGCCACTCATAGGCCACCTCTGTGTTTGAACACTCGGGGTCCAGAGCCCTGGGTGCT 3453
DB 1144 GCGCCACTCATAGGCCACCTCTGTGTTTGAACACTCGGGGTCCAGAGCCCTGGGTGCT 1203

```

```

QY 3454 ACTGCTAGCAGGGCTCTTCAATCAGACTCAGCTTGAACCTAGGTTCTTTTACCTGGG 3513
DB 1204 GTGGCAGCCGGGTGTCCACACAGACTCATCTGAAC----- 1241
QY 3514 GCTCTAGGAATTAGCCCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3573
DB 1242 -----TAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1289
QY 3574 CCGTTTCTCCCCCGCTGTGTGAGGGGCTCCACACAGACTTAAGCTAGGCACTATATGGACCA 3633
DB 1290 ---TTTCTCCCCCGCTGTGTGAGGGGCTCCACACAGACTTAAGCTAGGCACTATATGGACCA 1346
QY 3634 GCTGAATCCCATGTGTGAGGGG--CCATATAGACACAGGGGA--CTTGTCTTAGCTCACGTA 3691
DB 1347 GCGCAACCCCATGTGTGAGGGGCGCCATATGAGCAGGGGACCTTGTCTTAGCTAGGCA 1406
QY 3692 CCAGATGAGCTAAGTGATTAGGGGCTTGGATTACCGCCACTGTCTCCAGAGGCTATGGA 3751
DB 1407 CTTACGAGCTCAGTGAGAGGGGCTGTATTACCTCCACTGCCCCCAGGGGCTGTGGA 1466
QY 3752 TGAACCTGGTGGAG--CTGCCAGGCTTTTAC-----TGTTTTAACTTATTTTCACTGCTTT 3806
DB 1467 CAAACCGGCTGGGGGACTGCCCGAGGCTCCACCTGTTTATTAACTTATTTTCACTGCTTT 1526
QY 3807 ATAATAAGGAAACACTTAACGAAAAA 3838
DB 1527 ATAATAAGGAAACACTTAACGAAAAA 1558

RESULT 7
US-09-984-245-48
; Sequence 48, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095

```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-984-245-48

Query Match      13.7%; Score 528.4; DB 9; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;

QY 2995 CACATGAAGACTCACAGTGTGTGAAGAACCTTCGCTGTGCGCGTGTCCATACGCTCT 3054
DB 1 CACATGAAGACACACAGTGTGTGAAGAACCTTCGCTGTGCGCGTGTCTTATGCTCT 60

QY 3055 GCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCACACAGAGAGAAAGCCCTACAAG 3114
DB 61 CCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCATACAGAGAGAGCCCTACAAG 120

QY 3115 TGCCCCCTCTGTCCGATGCTGTGGCAACTCTGGCCAACTCAAGGCTCATGTGTCGATC 3174
DB 121 TGCCCCCTCTGTCCGATGCTGTGGCAACTCTGGCCAACTCAAGGCTCATGTGTCGATC 180

QY 3175 CACTCTGTGACAAACCTTTTCGCTGTAGCTTTGCACTACAGCTGCAACACAGAGATG 3234
DB 181 CACTCTGTGACAAACCTTTTCGCTGTAGCTTTGCACTACAGCTGCAACACAGAGATG 240

QY 3235 AACCTCAAAGCTCATATGCTGCGACACACGCGGAGAGAGCCCTTCGCTGTGCCACTGC 3294
DB 241 AACCTCAAAGCTCATATGCTGCGACACACGCGGAGAGAGCCCTTCGCTGTGCCACTGC 299

QY 3295 GCCTATACCAAGCCACTGGGCAACTACAGGCTCATCAGAGGTGCATGGCCATGGT 3354
DB 300 GCCTATACCAAGCCACTGGGCAACTACAGGCTCATCAGAGGTGCATGGCCATGGT 359

QY 3355 GGAGCAGAGAGGCTGTGTCTCTGCGCCCTGAGGCTGGGCCCCCACTCATAGCCCAACC 3414
DB 360 GGAGCAGAGAGGCTGTGTCTCTGCGCCCTGAGGCTGGGCCCCCACTCATAGCCCAACC 419

QY 3415 TCTGTTTGTAGACTCGGGTTCAGACGCTGGGTGTCTGCTGCTGTAGGAGGCTGTCTCAT 3474
DB 420 TCTGTTTGTAGACTCGGGTTCAGACGCTGGGTGTCTGCTGCTGTAGGAGGCTGTCTCAT 479

QY 3475 TCAGACTCACCTTGAATTAACCTAGGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
DB 480 ACAGACTCATCTGAATTAACCTAGGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 509

QY 3535 CCTCTGCAATTTTATAAATGAATAGACAGACCTTCCCTTTCTCCCGCTGTGCA 3594
DB 510 ---CCCATGTTTATACAGCGGACCAAGACCTT---TTTCTCCCGCTGTGCA 562

QY 3595 GGGGCTCCACACAGACTAACCTAGGCACTATATGACGAGCTCAATCCCATGTCAGGG 3654
DB 510 GGGGCTCCACACAGACTAACCTAGGCACTATATGACGAGCTCAATCCCATGTCAGGG 3654

```

```

RESULT 8
US-09-966-262-48
; Sequence 48, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30

```

```

Db 563 GGGGCTCCACACAGACTAAAGTACCTATAGGACCTATAGGACCCACCCCATCGGGGGGG 622
QY 3655 GG-CCATATAGACCCAGGGA-CTTGCTCTTAGCTCAGTACCCAGATGAGCTAAGTATTAG 3712
Db 623 GGCCCATATGGACCCAGGACCTTGCTTGAAGGCACTTCAAGAGCTCAGTGAGAAG 682
QY 3713 GGCCTTGGATTACCCGCACTGCTCCAGAGGCTATGATGAAGTCTGTTGGAG-CTGCC 3771
Db 683 GGGCCTGTATTACCTCCACTGCCCCCGGGGCTGTGGAACAAACCCGCTGGGGGACTGCC 742
QY 3772 CAGCCTTTTAC-----TGTTTAACTTATTTCACTGCTTTTATATAAGGAGAAACACTAACA 3827
Db 743 CAGCCTCCCACTGTTTATTTAACTTATTTCACTGCTTTTATATAAGGAGAAACACTAACA 802
QY 3828 GA 3829
Db 803 AA 804

```

Query Match 13.7%; Score 528.4; DB 10; Length 851;
Best Local Similarity 83.3%; Pred No 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;
QY 2995 CACATGAAGACTCACAGTGTGAGAAACCTTCGGTGTGCGCGTGTCCATACGCGCTCT 3054
Db 1 CACATGAAGACACACAGTGTGAGAACCTTCGGTGTGCGCGTGTCTTATGCGCTCT 60
QY 3055 GCTCCTCTGATACCTGAAACGGCCACGCGGTCCACAGAGAGAAAGCCCTACAAG 3114
Db 61 CCTCATCTGATACCTGAAACGGCCACGCGGTCCATACAGAGAGAGAGCCCTACAAG 120
QY 3115 TGCCCCCTCTGTCGCTATGCTGTGGCAACCTTGCCCAACCTCAAGCGTCAATGTCGCATC 3174
Db 121 TGCCCCCTCTGTCGCTATGCTGTGGCAACCTTGCCCAACCTCAAGCGTCAATGTCGCATC 180
QY 3175 CACTCTGGTACAAACCTTTTCGGTGTAGCTTTGCAACTACAGCTGCAACACGAGTATG 3234
Db 181 CACTCTGGTACAAACCTTTTCGGTGTAGCTTTGCAACTACAGCTGCAACACGAGTATG 240
QY 3235 AACCTCAACGTCATATGCTGCCACACACGCGGCGAGAACCTTCGCGTGTGCCACCTGC 3294
Db 241 AACCTCAACGTCATATGCTGCCACACACGCGGCGAGAACCTTCGCGTGTGCCACCTGC 299
QY 3295 GCCTATACCAAGCCACTGGGCAACTCAAGCGTCAAGAGTGTGATGCCATGT 3354
Db 300 GCCTATACCAAGCCACTGGGCAACTCAAGCGTCAAGAGTGTGATGCCATGT 359
QY 3355 GGAGCAGAGGCGCTGCTCTCTGCGCTGTGGGCTGGGCGCCACCTCATAGCCCAACC 3414
Db 360 GGAGCAGAGGCGCTGCTCTCTGCGCTGTGGGCTGGGCGCCACCTCATAGCCCAACC 419
QY 3415 TCTGTTTGTAGCACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGAGGGCTCTTCTAT 3474
Db 420 TCTGTTTGTAGCACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGAGGGCTCTTCTAT 479
QY 3475 TCAGACTCACTTGAACCTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
Db 480 ACAGACTCACTTGAACCTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 509
QY 3535 CCTCCTGCATTTTATACAAATGAACTAGAAACCACTTCCCTTTCTCCCGCTGCTCA 3594
Db 510 ---CCCATGTTTATACAGAGGCGCCAGAGCCACCTT---TTTCTCCCGCTGCTCA 562
QY 3595 GGGGCTCCACAGACTAACTAGGCACTATAGGACAGGCTGAATCCCATGTCAGGG 3654
Db 563 GGGGCTCCACAGACTAACTAGGCACTATAGGACAGGCTGAATCCCATGTCAGGG 622
QY 3655 GG-CCATATAGACAGAGGGA-CTTGTCTTAGCTCACGTACAGATGAGCTAAGTGAATTAG 3712
Db 623 GGCCCATATGACACAGGAGGACCTTTCCTTGTAGCTGAGGCACTTACAGAGCTCAGTGAGAAG 682
QY 3713 GGCCTTGGATTCACGGCACTGCTCCAGAGGCTATGAGAACTGTTGGAG-CTGCC 3771
Db 683 GGCCTTGGATTCACGGCACTGCTCCAGAGGCTATGAGAACTGTTGGAG-CTGCC 742
QY 3772 CAGCCTTTTAC----TGTTTTAACTTATTTTCACTGCTTTTATAATAAGGAAACACTAACA 3827

Db 743 CAGCCTCCACCTGTTTATTAACCTTATTTCACTGCTTTTATAATAAGGAAACACTAACA 802
QY 3828 GA 3829
Db 803 AA 804
RESULT 9
US-09-983-966-48
; Sequence 48, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48


```

RESULT 12
US-09-764-891-7626
; Sequence 7626, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7626
; LENGTH: 6271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7626

Query Match      11.6%; Score 444.8; DB 10; Length 6271;
Best Local Similarity 73.5%; Pred. No. 1e-120;
Matches 795; Conservative 0; Mismatches 227; Indels 60; Gaps 16;

QY 1 GGAGGTTGTAGACTTTTCAATCAAGATCAGTCAACGATGATGAAACGGGGAGCTTCG 60
DB 4950 GGAAGTTGTAGATTTCATTCATCAATCAAGATCAGCCAGCGTATGAAATGGGGAGCTTCG 5009

QY 61 GTTATTCTCATCTCTGGAAGAGGTGAGTCTAAAGTTGGGAGGAGGNGNT-----TG 115
DB 5010 GTTATTCTCATCTCTGGAAGAGGTGAGTCCAGGTTGGAGAGAGAGGGTGTCTGGT 5069

QY 116 GTTAAACAGANAATTTTTTTAGATGTGACAGAGGTGGAGTTAGATCACTGAA 175
DB 5070 CTGCACAGCCAGGGTTACCTTTTAACTCTGAGCTAAGTCGGAATCTGGGAGCATATTC 5129

QY 176 NACAAATAGT-----ACTTTGATTAAGANGSGTNTTCTTCATTTCCACG 224
DB 5130 AGTAATTTGTGACAGAGCTTAACCCCGACTCAAGTCCAAAGTCCACTTCTGTCAATTCCTCG 5189

QY 225 ATGPAGCC-TGGGCTTCAGAAATNACNGTTTTTTTTTTTGTGATAACTTGGCCCATTC 283
DB 5190 ACTTAGCCTTATGCTTATGACTAGCAGGTGGGCGGCTGTGTGATACCTCTTGGCCCAT 5249

QY 284 TGTCCTGCTTTCTGTATAGCAAGCAAGCTTCTTGTGTAATTT----- 333
DB 5250 CCTCTCTACACCTGACTCAGCAAG--CCAGCCCTCTTGTGGAGCTGTAGATTCCTGG 5307

QY 334 -----GGTTCAAAACCCCATGCTAACTGCTTAA-CTTTTTTGTAGCTCTGGATCAGT 387
DB 5308 CAAGGTGTTCAAAACCTGTGGCTGAGCTACCCCATCCCTTTCTTAGCTCTGATCAGT 5367

QY 388 GCTTGGCGCAGACACTTCTGGGATGGTACAGGGTGTGACATGACATGATCATCA 447
DB 5368 GCCTGGCACCAGACACTTCTGGGATGGTACAGGGTGTGACATGACATGATCATCA 5427

QY 448 TTTGTTTCAAGCCCGAAACACAGTACAGTCTCAGGACAGAGTGGCAAGAGAACTGG 507
DB 5428 TTTGTTTCAAGCCCGAAACACAGTACAGTCTCAGGACAGAGTGGCAAGAGAACTAG 5487

QY 508 AGGAGGCACTGTCCACGGAGGGGCTGAAGACACCGGCAACAGTGACAAA---AAGAAG 564
DB 5488 AGGAGGCTCTCTACTAGGGGGCTGAAGAAATGCAACAGCGACAGAGAGAAAGG 5547

QY 565 CCAAGAGGATTTAGTGTCTACCGGA-CCCTGCCCATGTGACTGTTTCTGAGCCCT-T 622
DB 5548 CCAAGCAGACTAGACTTATCCAGACCCCTGGCCCACTAGCTGTTTCTGAGCCCTCC 5607

QY 623 GGACCCGAGACTGATGTTTCTTGTCTTGTCTTGTAGCTTACGATGGGTATGAGGTGCA 682
DB 5608 GGACCTGAGACTGATGTTTCTTGTCTTGTCTTGTAGCTTACGATGGGTATGAGGTGCA 5667

QY 583 GGGG--GCTGGGTGCTTCTCAGCCCATTAACAAGAGGCCCCCCCCACCCCGCCACGC 740
DB 583 GGGG--GCTGGGTGCTTCTCAGCCCATTAACAAGAGGCCCCCCCCACCCCGCCACGC 740

```

```

RESULT 13
US-09-917-800A-636/c
; Sequence 636, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/232,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 636
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AI058436
US-09-917-800A-636

```

```

Query Match      10.4%; Score 399.2; DB 9; Length 496;
Best Local Similarity 93.0%; Pred. No. 8.5e-108;
Matches 452; Conservative 0; Mismatches 29; Indels 6; Gaps 3;

```


DATE TO BEAT: 4/24/1998

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 15:34:15 ; Search time 9218 Seconds
(without alignments)
12462.552 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 ggaaggttagactttatttc.....gaaaaa..... 3847

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1:	em_estba:**
2:	em_esthum:**
3:	em_estin:**
4:	em_estmd:**
5:	em_estov:**
6:	em_estpl:**
7:	em_estro:**
8:	em_htc:**
9:	gb_est1:**
10:	gb_est2:**
11:	gb_htc:**
12:	gb_est3:**
13:	gb_est4:**
14:	gb_est5:**
15:	em_estfun:**
16:	em_estcom:**
17:	em_gss_hum:**
18:	em_gss_inv:**
19:	em_gss_pln:**
20:	em_gss_vrt:**
21:	em_gss_fun:**
22:	em_gss_man:**
23:	em_gss_mus:**
24:	em_gss_pro:**
25:	em_gss_rod:**
26:	em_gss_phg:**
27:	em_gss_vrl:**
28:	gb_gss1:**
29:	gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	827.2	21.5	1247	11	AK052459 Mus muscu
2	732.4	19.0	906	9	AL534516 AL534516
3	716.2	18.6	735	13	BQ749274 UI-M-FD0-
4	687.2	17.9	990	12	BI668143 603295981

RESULT 1
AK052459

LOCUS

DEFINITION

AK052459 1247 bp mRNA linear HTC 20-SEP-2003

Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched

library, clone:D430028M17 product:similar to CDNA FLJ32203 FIS,

CLONE PLACE6003038, WEAKLY SIMILAR TO ZINC FINGER PROTEIN 84 [Homo

sapiens], full insert sequence.

ACCESSION AK052459.1 GI:26342688

VERSION HTC; CAP trapper.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, N., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

PUBMED 10349636

REFERENCE 2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, N., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

ALIGNMENTS

5	677	17.6	1069	13	BQ72985
6	673.4	17.5	1001	13	BQ056860
7	666.8	17.3	843	10	BF533875
8	644.2	16.7	929	13	BF533875
9	628	16.3	645	12	BI408818
10	623.6	16.2	694	10	AW930386
11	616.2	16.0	793	12	BG574616
12	613.8	16.0	628	10	AW536524
13	613.8	16.0	666	10	AW536842
14	613.8	16.0	692	12	BM114215
15	613.8	16.0	727	12	BM119847
16	612.8	15.9	616	12	BM245850
17	612.4	15.9	718	13	EX672987
18	607.4	15.8	652	12	BI903576
19	603.8	15.7	805	12	BI407148
20	602.8	15.7	629	10	AW539472
21	602.8	15.7	663	12	BI903756
22	587.8	15.3	632	14	CA450828
23	581.2	15.1	668	12	BG707470
24	576.8	15.0	581	13	C78804
25	567.2	14.7	703	9	AI327168
26	562.6	14.6	693	14	CB327866
27	561.2	14.6	566	10	AW538698
28	560.4	14.6	912	12	BG392489
29	559.4	14.5	905	9	AL565547
30	557.8	14.5	563	9	AI461955
31	557.6	14.5	939	10	BZ908881
32	556	14.5	710	9	AI385652
33	551.8	14.3	567	10	AW538638
34	547	14.2	568	12	BI439920
35	546.8	14.2	550	10	AW544493
36	546.8	14.2	615	10	BB484959
37	544	14.1	705	12	BI407530
38	540.8	14.1	558	13	BQ921766
39	536.8	14.0	653	12	BM936232
40	536	13.9	663	13	BQ195083
41	529	13.8	587	13	BQ747039
42	528.6	13.7	607	13	EX670592
43	526.8	13.7	733	13	BQ203743
44	526.8	13.7	736	13	BQ208560
45	525.2	13.7	692	12	BI285247

BQ072985	AGENCOURT
BQ056860	AGENCOURT
BF533875	602075222
BF533875	602075222
BI408818	602965769
AW930386	UFO6C03.X
BG574616	602596613
AW536524	G0105E04
AW536842	G0109D12
BM114215	L0801D02
BM119847	L0931C10
BM245850	K0730B02
EX672987	EX672987
BI903576	603168601
BI407148	602918774
AW539472	C0120C06
BI903756	603167805
CA450828	UI-M-FV0
BG707470	602672893
C78804	C78804 Mous
AI327168	mg47e10.X
CB327866	UI-R-F50
AW538698	C0110G03
BG392489	602410681
AL565547	AL565547
AI461955	UC43Q02.X
BZ908881	601497802
AI385652	me01G02.X
AW538638	C0110A08
BI439920	IC74C03.X
AW544493	C0182D09
BB484959	BB484959
BI407530	602918774
BQ921766	AGENCOURT
BM936232	UI-M-C00P
BQ195083	UI-R-CN1
BQ747039	UI-M-FA0
EX670592	EX670592
BQ203743	UI-R-DN1
BQ208560	UI-R-DY1
BI285247	UI-R-DB0

```

11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Osawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1247)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kigawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saichoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,I., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216]
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
1..1247
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:D430028M17"
/db_xref="MGI:2421883"
/db_xref="taxon:10090"
/clone="D430028M17"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="13 days embryo"
101..907
/note="unnamed protein product; putative
similar to CDNA FLJ32203 FIS, CLONE PLACE6003038, WEAKLY
SIMILAR TO ZINC FINGER PROTEIN 84 [Homo sapiens]
(SPTFLQ96ML3, evidence: PASTY, 75.2%ID, 58.9%length,
match=666)"
/codon_start=1

```

```

/protein_id="BAC35001.1"
/db_xref="GI:26342689"
/translatability="MGTAWEDSTSTSLPSLVHDESDTDEGPGALVLESDLLGLQDLBF
EEEEDEGDNHQLMGFERDSQGRAPGLPYGLSDSDSGGRLASAEVVEEP
ARGFEASGEXPPRCARCPYASALDNLKRVHVTGKPYKPLCTPYACNLANLKR
HGRTHSGKPPRCLNYSQNSMLKRLHMTGKPYKPCATCAITTTGHWNDYKRHQ
KVHGHGAGGPGRSAPGWAPPSPSVLSTGRPALCATGSRALHSDSP"
1227..1232
polyA_signal
/note="putative"
1247
polyA_site
/note="putative"
ORIGIN
Query Match 21.5%; Score 827.2; DB 11; Length 1247;
Best Local Similarity 88.8%; Pred. No. 2.9e-128;
Matches 895; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 2821 GGACAGGAACCTGGAGAGGCTGAGGGGAGCGAGCTGGGAGCTGCCATGTGTGGGCGCTGC 2880
DB 239 GAAGAGGAAGAGATGAAGGTGACGGCCACACACAGCAGCTCATGGGCTTTTGAGAGAGAC 298
QY 2881 ATGCCGAGGAGGCTGGAGGGGTTCCTCCTGGGGACCCAGGGGCTGGTGCACAAAGGC 2940
DB 299 TCTGAGGAGACTCTCAGGGGCCAGACTGACTTCCCTATGGGCTGAGTGACGACGAG 358
QY 2941 TTGGCTGTAGTTTATGCCCCCTTTGCCACTACTACCCCAACCACTGGCTCGGCACATG 3000
DB 359 TCTGGGGGGCGGGCGCACTAAGTGGGAGAGTGAAGTTGAGAAACAGCCAGCGGGTCCA 418
QY 3001 AAGACTCACAGTGGTGGAGAAACCTTCGCTGTGCGCTGTGCCATAGCCCTCTGCTCAT 3060
DB 419 GGGAGGCGCAGTGGTGAGAAACCTTCGCTGTGCGCTGTGCCATAGCCCTCTGCTCAT 478
QY 3061 CTGGATAACCTGAAACGGCACAGCGCGCTCCACACAGGAGAAAACCTTACAAGTGCCCC 3120
DB 479 CTGGATAACCTGAAACGGCACAGCGCGCTCCACACAGGAGAAAACCTTACAAGTGCCCC 538
QY 3121 CTCTGTCCGTATGCTGTGGCAACCTGCCACCTCAGCGTCATGTGGCATCCACTCT 3180
DB 539 CTCTGTCCGTATGCTGTGGCAACCTGCCACCTCAGCGTCATGTGGCATCCACTCT 598
QY 3181 GGTGACAAACCTTTTCGGTGTAGCTTTTGCAACTACAGCTGCAACACAGATGAAACCTC 3240
DB 599 GGTGACAAACCTTTTCGGTGTAGCTTTTGCAACTACAGCTGCAACACAGATGAAACCTC 658
QY 3241 AATCGTCATATGTCGGACACACAGGGCGAGAGCCCTTCGGCTGTGGCACCTGGCCCTAT 3300
DB 659 AATCGTCATATGTCGGACACACAGGGCGAGAGCCCTTCGGCTGTGGCACCTGGCCCTAT 718
QY 3301 ACCACAGGCACTGGGACAACTACAAGCGTCATCAGAGGTCATGGCCATGGTGAGCA 3360
DB 719 ACCACAGGCACTGGGACAACTACAAGCGTCATCAGAGGTCATGGCCATGGTGAGCA 778
QY 3361 GGAGGCGCTGCTCTCTGCTGCTGAGGGCTGGGCGCCACCTCATAGCCACCTCTGTT 3420
DB 779 GGAGGCGCTGCTGCTGCTGCTGAGGGCTGGGCGCCACCTCATAGCCACCTCTGTT 838
QY 3421 TTGAGCACTCGGGGTCACAGAGCCCTGGGTGCTACTGTGTAGCAGGGCTCTTCACTCAGAC 3480
DB 839 TTGAGCACTCGGGGTCACAGAGCCCTGGGTGCTACTGTGTAGCAGGGCTCTTCACTCAGAC 898
QY 3481 TCACCTTGAACCTAAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCTCT 3540
DB 899 TCACCTTGAACCTAAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCTCTCT 958
QY 3541 GCATTTTATACAACTGAACCTAGAAACCTTTTCCCTTTTCCCGCGCTGGTCAGGGGCT 3600
DB 959 GCATTTTATACAACTGAACCTAGAAACCTTTTCCCTTTTCCCGCGCTGGTCAGGGGCT 1018
QY 3601 CCACAGAGCTAACCTAGGCACTATATGACACAGCCTGAATCCCATGGTGAGGGGCTCAT 3660
DB 1019 CCACAGAGCTAACCTAGGCACTATATGACACAGCCTGAATCCCATGGTGAGGGGCTCAT 1078

```

```

QY 3661 ATAGACAGGGGACTTGTCTAGCTCACTACAGATGAGTAACTGATTAGGCGCTTGG 3720
Db 1079 ATAGACAGGGGACTTGTCTAGCTCACTACAGATGAGTAACTGATTAGGCGCTTGG 1138
QY 3721 ATTACCGGCACTGCTCCAGAGGCTATCGATGAACCTGTTGGAGCTGCCCCAGCCTTTT 3780
Db 1139 ATTACCGGCACTGCTCCAGAGGCTATCGATGAACCTGTTGGAGCTGCCCCAGCCTTTT 1198
QY 3781 ACTGTTTAACTTATTTCAGTCTGCTTATTAATAAGGAAACACTAACAG 3828
Db 1199 ACTGTTTAACTTATTTCAGTCTGCTTATTAATAAGGAAACACTAACAG 1246

RESULT 2
AL534516 906 bp mRNA linear EST 12-MAY-2003
LOCUS AL534516 Homo sapiens FETAL BRAIN Homo sapiens cDNA Clone
DEFINITION CS0DF004YCl2 5-PRIME, mRNA sequence.
ACCESSION AL534516
VERSION AL534516.2 GI:30540389
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12798009.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1341.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF004YB06QPI&cluster=1341.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DF004YB06QPI.
FEATURES
source
Location/Qualifiers
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF004YCl2"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dr) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 19.0%; Score 732.4; DB 9; Length 906;
Best Local Similarity 92.6%; Pred. No. 2.3e-112;
Matches 766; Conservative 2; Mismatches 59; Indels 0; Gaps 0;

QY 1917 AGTGGATCTCAAGATTCCTTCGACGAGTCTCTGGGCGCCCTGGTGTGGAGCGGATTT 1976
Db 80 AGTGGATCTCAAGATTCCTTCGACGAGGAGCCCGGCGCCCTGGTGTGGAGCGGATTT 139
QY 1977 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGATGAGTGCAGCCCA 2036
Db 140 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
QY 2037 CAACGACGAGTCTATGGGCTTTGAGAGAGTCTTGAAGGAGACTCTCAGGGGGCCAGACC 2096
Db 200 CAGTGACGAGTCTATGGGCTTTGAGAGAGTCTTGAAGGAGACTCTCTGGGGGGCCAGGCC 259

```

```

QY 2097 TGGACTTCCCTATGGGCTGAGTGACGACGAGTCTTGGGGGGCGCGCGCACTAAAGTGCGBA 2156
Db 260 TGGGCTTCCCTATGGGCTGAGCGACGATGATCTTGGGGGGCGCGCGCACTAAAGTGCGBA 319
QY 2157 GAGTGAAGTTAGAGAAACAGCCAGGAGGTCAGGGGAGCCAGGGGTGAGAGGCCAGAGCCC 2216
Db 320 GAGTGAAGTTAGAGAGACAGCCAGGAGGTCAGGGGAGCCAGGGGTGAGAGGCCAGAGCCC 379
QY 2217 AGCCTGTGAGTGTGGGGGGCCGACAGAGGTGAGGGGCGCTGTGTTGTGGGGCAGAGAGGCG 2276
Db 380 AGCCTGTGAGTGTGGGGGGCCGACAGAGGTGAGGGGCGCTGTGTTGTGGGGCAGAGAGGCG 439
QY 2277 GGGTGGGGGGCGCGCGCTGCCCCCACCGTTACTGTACTCATGCCGGCTGTGCGCTTTTGGT 2336
Db 440 GGGTGGGGGGCGCGCGCTGCCCCCACCGTTACTGTACTCATGCCGGCTGTGCGCTTTTGGT 499
QY 2337 GTCCCACTACTCGAGGCCACCTGAAGCGGCACATGACAGACACAGCCGGGAGAGAGCGCTT 2396
Db 500 GTCCCACTACTCGAGGCCACCTGAAGCGGCACATGACAGACACAGCCGGGAGAGAGCGCTT 559
QY 2397 CGCTGTGGCGCTGCCCCATACGNTACGCCAGTGTGTAACCTGACCGCCACATACCCG 2456
Db 560 CGCTGTGGCGCTGCCCCATACGNTACGCCAGTGTGTAACCTGACCGCCACATACCCG 619
QY 2457 CACCCATACTGGCAGAGAGCCCTACCGTTGTCCCACTGCCCTTTGCTCGCAGCAGCCT 2516
Db 620 CACCCATACTGGCAGAGAGCCCTACCGTTGTCCCACTGCCCTTTGCTCGCAGCAGCCT 679
QY 2517 GGGCAACCTGAGGGCGCATCAGCGCACCCACAGGGGCTCCCACTCTCCTGCTGCCAAC 2576
Db 680 GGGCAACCTGAGGGCGCATCAGCGCACCCACAGGGGCTCCCACTCTCCTGCTGCCAAC 739
QY 2577 CTGTGGCTTTCGATGCTGTCTCCAGCAGAACCGGCTCCAGTCCACAGACAGCAGCA 2636
Db 740 CTGTGGCTTTCGATGCTGTCTCCAGCAGAACCGGCTCCAGTCCACAGACAGCAGCA 799
QY 2637 GGGGCAATGCCCCGACGATCAGAAATGCGGTGATCTCTGCGCAGACTTGAGTCTTCATGT 2696
Db 800 GGGGCGGCTGCCCGGCGACCTGAGATGCTGTCTCTCTCCAGATTTGAGCCTCCATGT 859
QY 2697 GCACACAGTGTGTCAGCTTCTCTCCAGACTGTGGCAGCTGCGGG 2743
Db 860 GCCACAGGTGTGCCAGTTCTCTCCAGACTGTGGGAGTGGGG 906

RESULT 3
BQ749274 735 bp mRNA linear EST 17-JUL-2002
LOCUS BQ749274
DEFINITION UI-M-FD0-byf-1-07-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
IMAGE:5717526 5', mRNA sequence.
ACCESSION BQ749274.1 GI:21896061
VERSION BQ749274
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 735)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

Seq primer: pYX-5.	
Location/Qualifiers	
1..735	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6"	
/db_xref="taxon:10090"	
/clone="IMAGE:5717526"	
/tissue_type="whole brain"	
/dev_stage="embryo 12.5 dpc"	
/lab_host="DH10B (T1 phage resistant)"	
/clone_lib="NIH BMAP_FDO"	
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."	
ORIGIN	
Query Match	18.6%; Score 716.2; DB 13; Length 735;
Best Local Similarity	99.2%; Pred. No. 1.2e-109;
Matches 729; Conservative	0; Mismatches 5; Indels 1; Gaps 1;
QY	2838 GGCTGAGGGCAGCAGGCTGGAGCTGCCATGTGTGGCGCTGCATGCGAGGAGGCTGG 2897
Db	1 GGGTGAGGGCAGCAGGCTGGAGCTGCCATGTGTGGCGCTGCATGCGAGGAGGCTGG 60
QY	2898 AGGGGTGGCACT-GGGGGACCCGAGGCGCTGGTGACAAAGCTTGGCTGTAGTTAT 2956
Db	61 AGGGGTGGCACTGGGGGACCCGAGGCGCTGGTGACAAAGCTTGGCTGTAGTTAT 120
QY	2957 GCCCTTTGGCACTACTACCAACACCTGCTCGGCACATGAAGCTCAAGTGGTG 3016
Db	121 GCCCTTTGGCACTACTACCAACACCTGCTCGGCACATGAAGCTCAAGTGGTG 180
QY	3017 AGAAACCTTCCGCTGTGCGCGCTGTCCATACGCTCTGCTCATCTGGATACCTGAAC 3076
Db	181 AGAAACCTTCCGCTGTGCGCGCTGTCCATACGCTCTGCTCATCTGGATACCTGAAC 240
QY	3077 GGCAACGCGCTCCACAGAGGAGAAAGCCCTACAGTGCCTCCCTCTGCTGATGCTT 3136
Db	241 GGCAACGCGCTCCACAGAGGAGAAAGCCCTACAGTGCCTCCCTCTGCTGATGCTT 300
QY	3137 GTGGCACTGGCCAACTCAAGCGTCACTGGTGGCATCCACTGTGGTGACAACTTTTC 3196
Db	301 GTGGCACTGGCCAACTCAAGCGTCACTGGTGGCATCCACTGTGGTGACAACTTTTC 360
QY	3197 GTGTAGCTTTGCACTACAGTGGCAACAGAGTATGAACTCAAGCTCATATGCTGTC 3256
Db	361 GTGTAGCTTTGCACTACAGTGGCAACAGAGTATGAACTCAAGCTCATATGCTGTC 420
QY	3257 GACACAGGGGAGAGCCCTTCGCTGTGCGCACCTCGGCTATACACAGGCCACTGGG 3316
Db	421 GACACAGGGGAGAGCCCTTCGCTGTGCGCACCTCGGCTATACACAGGCCACTGGG 480
QY	3317 ACAACTCAAGCGTATCAGAGGTGATGCGCATTGTGTGAGCAGGAGGCGCTGTCT 3376
Db	481 ACAACTCAAGCGTATCAGAGGTGATGCGCATTGTGTGAGCAGGAGGCGCTGTCT 540
QY	3377 CTGCCCCTGAGGGCTGGGCGCCCACTCATAGCCCAACCTCTGTTTTCAGCACTCGGGGTC 3436
Db	541 CTGCCCCTGAGGGCTGGGCGCCCACTCATAGCCCAACCTCTGTTTTCAGCACTCGGGGTC 600

Seq primer: pYX-5.	
Location/Qualifiers	
1..735	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6"	
/db_xref="taxon:10090"	
/clone="IMAGE:5717526"	
/tissue_type="whole brain"	
/dev_stage="embryo 12.5 dpc"	
/lab_host="DH10B (T1 phage resistant)"	
/clone_lib="NIH BMAP_FDO"	
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."	
ORIGIN	
Query Match	17.9%; Score 687.2; DB 12; Length 990;
Best Local Similarity	85.6%; Pred. No. 7.7e-105;
Matches 847; Conservative	0; Mismatches 129; Indels 13; Gaps 7;
QY	2352 CCACTGAAGCGGCATGACACACAGCGGGAGAGCGTTCCGCTGTGGCGCTG 2411
Db	1 CCACTGAAGCGGCATGACACACAGCGGGAGAGCGTTCCGCTGTGGCGCTG 60
QY	2412 CCATACGCTCAGCCCACTGGTCAACTGACGGGACATACCCCACTACTGGCGA 2471
Db	61 CCCTAGCGCTCAGCCCACTGGTCAACTGACGGGACATACCCCACTACTGGCGA 120

REFERENCE 1 (bases 1 to 1069)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaapb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12799 row: d column: 19
High quality sequence stop: 671.

FEATURES
Location/Qualifiers
1..1069
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5757282"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 17.6%; Score 677; DB 13; Length 1069;
Best Local Similarity 92.1%; Pred. No. 3.8e-103;
Matches 735; Conservative 0; Mismatches 61; Indels 2; Gaps 2;
2326 TGCGCTTTCGTTGCCACTACTCGAGCCACTGAGCGGCACATGACAGACACAGAGGGG 2385
Db 1 TGCACCTTCGTGTCCTACTCTCGAGCCACTGAGCGGCACATGACAGACACAGAGGGG 60
2386 GAGAACGCGTTCCGCTGTGGCGCTGCCATACGCTCAGCCAGTTCGTCACCTGACG 2445
Db 61 GAGAACGCGTTCCGCTGTGGCGCTGCCATACGCTCAGCCAGTTCGTCACCTGACG 120
2446 CGACATACCCGACCCACATCTGGCGAGAGCCCTACCGTTGTCCGCTGCGCCCTTTGCC 2505
Db 121 CGACATACCCGACCCACATCTGGCGAGAGCCCTACCGTTGTCCGCTGCGCCCTTTGCC 180
2506 TGCAGCAGCTGGGCAACCTGAGCGGCATCAGCGCCACCCACAGAGGCTCCCACTCCT 2565
Db 181 TGCAGCAGCTGGGCAACCTGAGCGGCATCAGCGCCACCCACAGAGGCTCCCACTCCT 240
2566 CCCTGCCCAACCTGTGGCTTTTCGATGTGTGCTTCAGACCAACCCGGGCTCCAGTCCC 2625
Db 241 CCCTGCCCGACCTGTGGCTTCGCTGTGTACTCCAGCACCGCCGCGCTCCAGTCCC 300
2626 ACAGAGCAGGAGGGGACATGCCCGGAGTACAGAAATGCGTGTCTCCGAGACTG 2685
Db 301 ACAGAGCAGGAGGGGCGGTGCCCGGAGTACAGAAATGCGTGTCTCCGAGACTG 360
2686 AGCTTTTCATGTGCCACCAAGGTGGTCCAGTTTCTTCGACAGTGTGGGAGCTCGGGGT 2745
Db 361 AGCTTCATGTGCCACCAAGGTGGTCCAGTTTCTTCGACAGTGTGGGAGCTCGGGGT 420
2746 GAAGGGGAGAGCTTGTGTGGAACTGGATCCGAAACCACTGCCAGAGCTACTGTCCCTGG 2805
Db 421 GAAGGGGAGAGGCTTGTGTGGAACTGGATCCGAAACCACTGCCAGAGCTACTGTCCCTGG 480
2806 ACCTGCCGGGGGTGTGGACAGGAACCTGAGGAGGGGTGAGGGGAGCTGGAGCTGCC 2865

2472 GAAGCCCTACGTTGTCCCACTGCCCCCTTTCCTGACAGAGCTGGGCACTGAGCG 2531
Db 121 GAAGCCCTACGTTGTCCCACTGCCCCCTTTCCTGACAGAGCTGGGCACTGAGCG 180
2532 GCATCAGCGCACCCACACAGAGGCTCCCACTCCCTGCTGCCCAACCTGTGTGCTTCGATG 2591
Db 181 GCATCAGCGTACCCACAGAGGCCCCCACTCCCTGCTGCCCAACCTGTGTGCTTCGCTG 240
2592 CTGTGCTCCAGACCAACCCGGGCTCCAGTCCCAACAGAGAGAGGGGACAAATGCCCG 2651
Db 241 CTGTACTCCAGACCAAGCCGCGCTCCAGTCCCAACAGAGAGAGGGGCGGGTGCCTCC 300
2652 AGCATCAGAAATGCGCTGATCCTGCCAGACTTGTAGTCTTCATGTGCCACAGTGTGTG 2711
Db 301 GCGACTGAAGATGCTGTCTCTCTCCAGATTTGAGCTTCCATGTGCCACAGTGTGTG 360
2712 CAGTTTCTCTGCCAGACTGTGGGACGCTCGGGGTGAAGGGGAGAGCTTGTGTGAATCG 2771
Db 361 CAGTTTCTCTGCCAGACTGTGGGACGCTCGGGGTGAAGGGGAGAGCTTGTGTGAATCG 420
2772 ATCCGAACCACTGCCAGAGTACTGTTCCTTGGACCTGCGGGGCTGTGGACAGGAAC 2831
Db 421 ATCCGAACCACTGCCAGAGTACTGTTCCTTGGACCTGCGGGGCTGTGGACAGGAAC 480
2832 CGAGAGGCTGAGGGCAGCAGGCTGGAGCTGCCATGTGTGGCGCTGCATGCCAGGAGA 2891
Db 481 CGAGAGGCTGAGGGTGTAGTGTGGCTGGAGCTGCCATGTGTGGCGCTGCATGCCAGGAGA 540
2892 GCGTGAAGGGGTGCACGTGGGGGACCCAGAGGCTGTGTGACAAAGCTTTGCTGTAG 2951
Db 541 GCGTGAAGGGGTGCACGTGGGGGACCCAGAGGCTGTGTGACAAAGCTTTGCTGTAG 600
2952 TTTATGCCCTTTGCCACTCACTACCCCAACCACTGGCTCGGCGCATGAAGACTCAAG 3011
Db 601 CCTCTG-CCCTTTGCCACTCACTATCCCAACCACTGGACCGGCACATGAAGACACAG 659
3012 TGTGAGAAACCTTCCGCTGTGCGCTGTGCTATGAGCTTGTGCTATCTGTGTAACCT 3071
Db 660 TGTGAGAAACCTTCCGCTGTGCGCTGTGCTATGAGCTTGTGCTATCTGTGTAACCT 719
3072 GAAACGGCACAGCGGTGCACACAGAGAGAGCCCTTCAAGTGGCCCTCTGTCCGTA 3131
Db 720 GAAACGGCACAGCGGTGCACACAGAGAGAGCGCTTAAAGTGGCCCTCTGACCTT 779
3132 TGCC-----TGTGGCAACCTGGCAACCTCAAGGCTCATGGTGCATCTCT---GGTG 3184
Db 780 ATGCGCTGTGGCAAACTTTGGCCAACTCAAGGCTCAATGGTCCGAATCACCTCTGGGTG 839
3185 ACAACCTTTTCGCTGTAGCTTT--GCACTACAGTGCACACAGAGTATGAACCTCAA 3242
Db 840 ACAACCTTTTCGAGTGTAGCTTTTGGGAATAACAGTTTGAACAGAGAGTAACTCA 899
3243 ACCTCATATGCTGCGACAC-ACGGGGAGAGGCTTTCCGCTGTGCCACC-TGCGCTAT 3300
Db 900 AAGTACATCTGCAGCAACACAGCGAGAAACCTTACCGTGTGCCACCTTGGCATAT 959
3301 AC-CACAGGCACTGGAGCACTACAGC 3328
Db 960 ACACAGGGCACTGGAGCACTTAAAGGC 988

RESULT 5
BQ072985 1069 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_6817897 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5757282
DEFINITION 5', mRNA sequence.
ACCESSION BQ072985
VERSION BQ072985.1 GI:19902031
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

481 ACCTGCGCGGCTGTGGACAAGAGCTGGAGGGTGAAGGCTAGTGGCTGGGAGCTGCC 540
2866 ATGTGTGGGCGCTGATCGGAGAGAGGCTGGAGGGGTTGGCACTGGGGGACCCAGGGC 2925
541 ATGTGTGGGCGCTGATCGGAGAGAGGCTGGAGGGGTTGGCACTGGGGGACCCAGGGC 500
2926 CTGTGTGACAAAGGCTTTGGCTGTAGTTTATGCTCCCTTTGGCACTCACTACCCCAACCA- 2984
601 CCCAGTGACAAAGGCTTTGGCTGTAGCTCTGGCTCTGGCTTTGGCACTCACTATCCCAACCA 660
2985 CTGTGTGCGGCAATGAGAGCTCACTGTGTGTGAGAAACCTTCCGCTGCGCGCTGTCC 3044
661 CTGTGTGCGGCGCATGAGAGCACTGAGTGTGTGAGAAACCTTCCGCTGCGCGCTGTCC 720
3045 ATAGCGCTCTGCTCATCTGATTAACCTGAAACCGGACCGGCTGCCACACAGGA-GAAA 3103
721 TTATGCTCTGCTCATCTGATTAACCTGAAACCGGACCGGCTGCCACAGGAGGA 780
3104 AGCCCTACAAGTGCCCCC 3121
781 AGCCCTACCAATGCCCC 798

RESULT 6
BQ056860 1001 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_6940531 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5809303
5', mRNA sequence.
ACCESSION BQ056860
VERSION BQ056860.1 GI:19816200
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NTH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2054 row: 1 column: 08
High quality sequence stop: 576.
FEATURES
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5809303"
/tissue type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone lib="NTH_MGC_99"
note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 17.5%; Score 673.4; DB 13; Length 1001;
Best Local Similarity 91.8%; Pred. No. 1.5e-102;

```

```

Matches 765; Conservative 0; Mismatches 63; Indels 5; Gaps 5;
QY 2340 CCATCTACTCGAGCGACCTGAAGCGGCGCATGTCGAGACACACAGCGGGGAGAAAGCGTTCCG 2399
Db 5 CCATCTACTCGAGCGACCTGAAGCGGCGCATGTCGAGACACACAGCGGGGAGAAAGCGTTCCG 64
QY 2400 CTGTGGCGGCTGCCATATACGCTCAGCCCGCTGCTCAACTGACGGGACATACCCGCGAC 2459
Db 65 CTGTGGCGGCTGCCCGCTTACGCTCAGCCCGCTGCTCAACTGACGGGACATATACCCGCGAC 124
QY 2460 CCATCTACTGGCGAGAGCCCTTACCGTTGTCCCGCACTGCGCCCTTTGCTGCTGCGAGCGCTGGG 2519
Db 125 CCACACTGGCGAGAGCCCTTACCGCTGTCCCGCACTGCGCCCTTTGCTGCTGCGAGCGCTGGG 184
QY 2520 CAACCTGAGGGGCGATCAGCGCCACACAGGGGCTCCCACTCTCTCCCTGCGCAACCTG 2579
Db 185 CAACCTGAGGGGCGATCAGCGCTACCCACGAGGGGCCCCCACTCTCTCCCTGCGCAACCTG 244
QY 2580 TGGCTTTTCGATGCTGTGCTCCACGACCAACCCCGGCTCCCACTGTCACAGAGCGAGGG 2639
Db 245 TGGCTTTTCGCTGTGCTTCCACGACCGCGGCTCCCACTGTCACAGAGCGAGGG 304
QY 2640 GACAATGCCCCGACGATCAGAAATGCGCTGATCTCTGCGAGACTTGAGTCTTCTATGTGCC 2699
Db 305 GCGGTGCCCCGCGGACCTGAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 364
QY 2700 ACCAGTGTGTCGCTGCTTTCCTGCGCAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGAGCTT 2759
Db 365 ACCAGTGTGTCGCTGCTTTCCTGCGCAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGGGCT 424
QY 2760 GTGTGGAACTCGATCCGAAACCACTGCGCAGACTACTGTTCCCTTTGGAGCTGCCGGGGCTG 2819
Db 425 GTGTGGAACTCGATCCGAAACCACTGCGCAGACTACTGTTCCCTTTGGAGCTGCCGGGGCTG 484
QY 2820 TGGACAGAACTGGAGGAGGCTGAGGGGCGAGAGGCTGCGAGCTGCCAGTGTGTGGGCGCTG 2879
Db 485 TGGACAGAACTGGAGGAGGCTGAGGGGCTGAGGGTAGTGGCTGGAGAGTCCATGTGTGGGCGCTG 544
QY 2880 CATGCGAGGAGAGGCTGGAGGGGTTG-CACTGGGGGACCCCGAGGGCTGCTGGTACAAAG 2938
Db 545 CATGCGAGGAGAGGCTGGAGGGGTTG-CACTGGGGGACCCCGAGGGCTGCTGGTACAAAG 604
QY 2939 GCTTTGCTGTAGTTTATGCGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2998
Db 605 GCTTTGCTGTAGTTTATGCGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
QY 2999 TGAAGACTCACA-GTGGTGAGAAACCTTCCGCTGTGCGGCTGTGCGGCTGTGCTGCTGCTGCT 3057
Db 665 TGAAGACTCACA-GTGGTGAGAAACCTTCCGCTGTGCGGCTGTGCGGCTGTGCTGCTGCTGCT 724
QY 3058 CATCTGGATTAACCTGAAACGGGACAGGGGCTGCCACACA-GGAGAAAGCCCTTACAAGT- 3115
Db 725 CATCTGGATTAACCTGAAACGGGACAGGGGCTGCCACACA-GGAGAAAGCCCTTACAAGT 784
QY 3116 GCCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3167
Db 785 GCCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837

RESULT 7
BQ533875 843 bp mRNA linear EST 11-DEC-2000
LOCUS 602075222F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:421245 5',
DEFINITION mRNA sequence.
ACCESSION BQ533875
VERSION BQ533875.1 GI:11621238
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 843)
NIH-MGC http://mgc.nci.nih.gov/.

```

Db	601	GGTGCAAGGCTTTGCGCTGTAGTTATGCGCCCTTGTGCCACTCACTACCCCAAC--ACC	658
Qy	2988	GGCTGCGCATGAAAGACTACAGTGGTGAGAAACCTTTCGGCTGTGCGCGCTGTCCATA	3047
Db	659	TGGTGGGCACATGAAGACTCACAGTGGTGAGAAACCTTTC---GTGTGCGCGCTGTCCATA	715
Qy	3048	CGCTCTGCTCATCTGATTAACCTGAACCGGACCAGCGGTGCCACACAGGAGAAAGGCC	3107
Db	716	CGCTCTGTG-TCACTGTGATTAACCTGAAG-GGACCATTGCGT-CACACAGGAGAAAGGCC	772
Qy	3108	CTACAAGTGCCCC	3120
Db	773	TACCAAGTGCCCC	785
RESULT 8			
BX453491/c			
LOCUS			
DEFINITION			
BX453491 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone			
CS0DF004YC12 3-PRIME, mRNA sequence.			
BX453491			
ACCESSION			
BX453491.1 GI:31020508			
VERSION			
EST.			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 929)			
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 1341.r For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0BAI009ZF12_CS00862_2&cluster=1341.r.			
Contact : Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Paraday Avenue Genoscope sequence ID : CS0BAI009ZF12_CS00862_2.			
Location/Qualifiers			
1..929			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon.9606"			
/clone="CS0DF004YC12"			
/tissue_type="FETAL BRAIN"			
/dev_stage="fetal"			
/clone_lib="Homo sapiens FETAL BRAIN"			
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA			
was primed with a NotI-oligo(dT) primer. Five prime end			
enriched, double-strand cDNA was digested with Not I and			
cloned into the Not I and EcoRV sites of the pCMVSPORT 6			
vector. Library was not normalized."			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Query			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			

QY 3666 CCAGGGGACCTTGTCTTAGCTCAGCTACCATGATGAGCTAAGTATGAGGCTTGGATTCA 3725
 Db 107 CCAGGGGACCTTGTCTTAGCTCAGCTACCATGATGAGCTAAGTATGAGGCTTGGATTCA 48
 QY 3726 CCAGGACCTGCTCCAGAGGCTATGGATGAAGTCTGTTGGGA 3765
 Db 47 CCAGGACCTGCTCCAGAGGCTATGGATGAAGTCTGTTGGGA 8

RESULT 10
 AW990386/c
 LOCUS AW990386 694 bp mRNA linear EST 02-JUN-2000
 DEFINITION u06c03.x1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:1510564 3', mRNA sequence.
 ACCESSION AW990386
 VERSION AW990386.1 GI:8186118
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 694)
 REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Tumor Gene Index
 COMMENT
 Other_ESTs: uf06c03.y1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:937416
 High quality sequence stop: 439.
 Location/Qualifiers
 1..694
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:1510564"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /clone_lib="Soares mammary_gland_NMLMG"
 /note="Vector: p77T3-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified p77T3 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source
 Query Match 16.2%, Score 623.6; DB 10; Length 694;
 Best Local Similarity 94.8%; Pred. No. 3.7e-94;
 Matches 655; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
 QY 3139 GGCAACCTGGCCAACTCAAGCGTCATGGTGGCATCCACTGTGTGACAAACCTTTTCGG 3198
 Db 690 GGCAACCTGGCCAACTCAAGCGGTGATGGTGGCATCCACTGTGTGACCACTTTTCGG 631
 QY 3199 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTCGGA 3258
 Db 630 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTGC-A 572
 QY 3259 CACAGGGCGAGAGCCCTTCGGTGTGCCACCTCGCCTATACACAGGCCCACTGGGAC 3318
 Db 571 CACCGGGCGAGAGCCTTCGGTGTGCCACCTCGCCTATCCCTAGGGCCATTGGGAC 512
 QY 3319 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 3378
 Db 511 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 452

ORIGIN

Query Match 16.2%, Score 623.6; DB 10; Length 694;
 Best Local Similarity 94.8%; Pred. No. 3.7e-94;
 Matches 655; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
 QY 3139 GGCAACCTGGCCAACTCAAGCGTCATGGTGGCATCCACTGTGTGACAAACCTTTTCGG 3198
 Db 690 GGCAACCTGGCCAACTCAAGCGGTGATGGTGGCATCCACTGTGTGACCACTTTTCGG 631
 QY 3199 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTCGGA 3258
 Db 630 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTGC-A 572
 QY 3259 CACAGGGCGAGAGCCCTTCGGTGTGCCACCTCGCCTATACACAGGCCCACTGGGAC 3318
 Db 571 CACCGGGCGAGAGCCTTCGGTGTGCCACCTCGCCTATCCCTAGGGCCATTGGGAC 512
 QY 3319 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 3378
 Db 511 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 452

QY 3379 GCCCTGAGGCTGGGCCCCCAGCTCATAGCCAGCTCTGTTTTGAGCACTCGGGTCCA 3438
 Db 451 GCCCTGAGGCTGGGCCCCCAGCTCATAGCCAGCTCTGTTTTGAGCACTCGGGTCCA 392
 QY 3439 GCAGCCCTGGGCTGCTACTGTGTAGCAGGCTCTTCACTCAGACTCACCTTGAACCTAG 3498
 Db 391 GCAGCCCTGGGCTGCTACTGTGTAGCAGGCTCTTCACTCAGACTCACCTTGAACCTAG 332
 QY 3499 GTTCTTTTACCTGGGCTCTAGGAATAGCCTATGCTCTCTGCTGATTTTACAAATGAA 3558
 Db 331 GTTCTTTTACCTGGGCTCTAGGAATAGCCTATGCTCTCTGCTGATTTTACAAATGAA 272
 QY 3559 CTAGAAACCACTTTCCCTTTCTCCCCCGCTGGTCCAGGGCTCCACACAGACTAACCTAG 3618
 Db 271 CTAGAAACCACTTTCCCTTTCTCCCCCGCTGGTCCAGGGCTCCACACAGACTAACCTAG 212
 QY 3619 GCATATATGACACAGCCTGAAATCCATGGTCAGGGGCCATATAGACAGGGGACTTGT 3678
 Db 211 GCATATATGACACAGCCTGAAATCCATGGTCAGGGGCCATATAGACAGGGGACTTGT 152
 QY 3679 CTTAGCTCAGTACAGATGAGCTAAGTATTAGGGCTTGGATTCAACCGGCACTGCTCC 3738
 Db 151 CTTAGCTCAGTACAGATGAGCTAAGTATTAGGGCTTGGATTCAACCGGCACTGCTCC 92
 QY 3739 CAGAGCTATGATGACACTGGTTGGAGCTGCCAGCCTTTTACTGTTTAACTTATTC 3798
 Db 91 CAGAGCTATGATGACACTGGTTGGAGCTGCCAGCCTTTTACTGTTTAACTTATTC 32
 QY 3799 AGTGTCTTATATAAGGAAACACTAAACAGA 3829
 Db 31 AGTGTCTTATATAAGGAAACACTAAACAGA 1

RESULT 11
 BG574616
 LOCUS BG574616 793 bp mRNA linear EST:10-APR-2001
 DEFINITION 602596613F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705610 5',
 mRNA sequence.
 ACCESSION BG574616
 VERSION BG574616.1 GI:13582269
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 793)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10572 row: i column: 03
 High quality sequence stop: 750.
 Location/Qualifiers
 1..793
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4705610"
 /issue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for

FEATURES

source
 Query Match 16.2%, Score 623.6; DB 10; Length 694;
 Best Local Similarity 94.8%; Pred. No. 3.7e-94;
 Matches 655; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
 QY 3139 GGCAACCTGGCCAACTCAAGCGTCATGGTGGCATCCACTGTGTGACAAACCTTTTCGG 3198
 Db 690 GGCAACCTGGCCAACTCAAGCGGTGATGGTGGCATCCACTGTGTGACCACTTTTCGG 631
 QY 3199 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTCGGA 3258
 Db 630 TGAGCCTTTCAGTACAGCTGCACAGGATGATGACCTCAACCTCATATGCTGC-A 572
 QY 3259 CACAGGGCGAGAGCCCTTCGGTGTGCCACCTCGCCTATACACAGGCCCACTGGGAC 3318
 Db 571 CACCGGGCGAGAGCCTTCGGTGTGCCACCTCGCCTATCCCTAGGGCCATTGGGAC 512
 QY 3319 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 3378
 Db 511 AACTACAGCGTCATCAGAGGTGATGCCATGGTGGAGCAGAGGCGCTGCTCTCT 452

full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 16.0%; Score 616.2; DB 12; Length 793;
Best Local Similarity 89.9%; Pred. No. 66-93; 74; Indels 6; Gaps 5;
Matches 716; Conservative 0; Mismatches 74; Indels 6; Gaps 5;

QY 2137 GCGCGGCTAAAGTGGCGAGAGTGAAGTTGAGGAACACGAGCGGTGCCAGGGAGGCC 2196
DB 1 GCGCGGCTAAAGTGGCGAGAGTGAAGTTGAGGAACACGAGCGGTGCCAGGGAGGCC 58

QY 2197 AGGGGTGAGAGCGCCAGCCCTGTCAGCTGTGTGGGGGCGGACAGGTGAGGGCGC 2256
DB 59 AGGTGTGAGAGCGCCAGCCCTGTCAGCTGTGTGGGGGCGGACAGGTGAGGGCGC 118

QY 2257 TGTGTGGGAGAGGG 2316
DB 119 TGTGTGGGAGAGGG 178

QY 2317 TCCCGGCTGTGC-GCTTTCGTGTCCACTACTCTGAGAGCCACTGAGAGCGGCACATGCAGAC 2375
DB 179 TCCCGGCTGTGC-GCTTTCGTGTCCACTACTCTGAGAGCCACTGAGAGCGGCACATGCAGAC 238

QY 2376 ACACAGCGGGAGAGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 2435
DB 239 ACACAGCGGGAGAGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 298

QY 2436 CAACCTGAGCGGACATACCCGACCCACATCTGGCGAGAGCCCTACCGTGTCCCACTG 2495
DB 299 CAACCTGAGCGGACATACCCGACCCACATCTGGCGAGAGCCCTACCGTGTCCCACTG 358

QY 2496 CCGCTTGTGCTGAGAGAGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 2555
DB 359 CCGCTTGTGCTGAGAGAGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 418

QY 2556 TCCACTCTCTCCCTGCGGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 2615
DB 419 CCGACTCTCTCTCCCTGCGGCGGCTTCCGCTGTGGCGGCTGCCATACGCTCAGCCAGTTCGT 478

QY 2616 TCCAGTTCGCGAG 2675
DB 479 TCCAGTTCGCGAG 538

QY 2676 GCGAGA-CTTGAGCTTCATGTCACACAGGTGGTCCAGTTCCTGCCAGACTGTGGGC 2734
DB 539 TCCAGATTTGAGCTTCATGTCACACAGGTGGTCCAGTTCCTGCCAGACTGTGGGC 598

QY 2735 AGCTGGGGGTGAAGGGGAGAGCTTGTGTGAATCTGGATCCGAAACCACTGCCAGAGCTAC 2794
DB 599 AGCTGGGGGTGAAGGGGAGAGGCTTGTGTGAATCTGGATCCGAAACCACTGCCAGAGCTAC 657

QY 2795 TGTTCCTTGGAGCTGCGGGGGCTG-TGGACAGAACTGGAGAGGGGTGAGGSCAGCAGG 2853
DB 658 TATTCCTTGGAGCTGCGGGGGCTG-TGGACAGAACTGGAGAGGGGTGAGGSGAGTGTGG 717

QY 2854 CTGGAGCTGCCATGTGTGGCGCTGCATCCGAGGAGAGCTGGAGGGGTGCCACTGGG 2913
DB 718 CTGGAGCTGCCATGTGTGGCGCTGCATCCGAGGAGAGCTGGAGGGGTGCCACTGGG 777

QY 2914 GGACCCAGGGCCCTG 2929
DB 778 GGGCCCCAGGCCCCAG 793

RESULT=12
AW536524/c 628 bp mRNA linear EST 31-AUG-2000
LOCUS AW536524.1
DEFINITION Mus musculus E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0105E04 3', mRNA sequence.
ACCESSION AW536524
VERSION AW536524.1 GI:7178941
KEYWORDS EST.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 628)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagataja,R., Doi,H., Wood,W.H., III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
PUBMED 10922068
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@leuon.grc.nia.nih.gov
Plate: G0105 row: E Column: 04
Seq primer: -21M13 Forward
High quality sequence stop: 628
POLYA-Yes

Location/Qualifiers
1..628
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nia:EST:G0105E04-3"
/db_xref="taxon:10090"
/clone="G0105E04"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/notes="Vector: pSPORT1 (Gibco/BRL Life Technology); Site 1: SalI; Site 2: NotI; Total RNAs were extracted from 6 Embryo. The double-stranded cDNA was synthesized by Gibco's kit with an oligo(dT) primer [NotI primer-adaptor from GibcoBRL] [5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 0.51ug of mRNA. The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker JJ-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko."

ORIGIN

Query Match 16.0%; Score 613.8; DB 10; Length 628;
Best Local Similarity 99.7%; Pred. No. 1.7e-92; Indels 0; Gaps 0;
Matches 615; Conservative 0; Mismatches 2;

QY 373 AGCTGCTGGATCAGTTCCTGGCGCCAGACACTTCTGGGATGGTACAGGGTGTGACACA 432
DB 617 AGCTGCTGGATCAGTTCCTGGCGCCAGACACTTCTGGGATGGTACAGGGTGTGACACA 558

QY 433 TGACGTGATCATCATTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGCAGAGAGTG 492
DB 557 TGACGTGATCATCATTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGCAGAGAGTG 498

QY 493 GCAAGAGGAACTGGAGGAGGACCTGTCCACGAGGGGGGCTGAAAGACACCCGGCAACAGTG 552
DB 497 GCAAGAGGAACTGGAGGAGGACCTGTCCACGAGGGGGGCTGAAAGACACCCGGCAACAGTG 438

Db	497	CGAAGAGAAACTGAGAGGCACTGTCACAGGAGGGGCTGAAGACACCGGCACAGTG	433		
QY	553	ACAAAGAGAGGCCCAAGAGGGATTAGTGGTCAACCGAGACCTGCCCATGTGAAGTCTTTT	612		
Db	437	ACAAAGAGAGGCCCAAGAGGGACTAGTGGTCAACCGAGACCTGCCCATGTGAAGTCTTTT	378		
QY	613	CTGAGCCCTTGGACCCGAGACTGAGTCTTCTGCTGTCCTTTAGCCTTAGCAGTGGGTAT	672		
Db	377	CTGAGCCCTTGGACCCGAGACTGAGTCTTCTGCTGTCCTTTAGCCTTAGCAGTGGGTAT	318		
QY	673	GAGGTGTGCAGGGGGCTGGTGGCTTCTCAGGCCATTACAAGAGGGCCCCCACCC	732		
Db	317	GAGGTGTGCAGGGGGCTGGTGGCTTCTCAGGCCATTACAAGAGGGCCCCCACCC	258		
QY	733	CCCCACGGGCGAGCCTGGGAGGCTCTGCTGCTCTTAAAGCCTCCTTACTCTCTTGGGC	792		
Db	257	CCCCACGGGCGAGCCTGGGAGGCTCTGCTGCTCTTAAAGCCTCCTTACTCTCTTGGGC	198		
QY	793	TCATCGACTATCGGTTCTGTGCCCTGTGCTCTGTTGTCTGGAGGGAAGGACTGCTAGTTC	852		
Db	197	TCATCGACTATCGGTTCTGTGCCCTGTGCTCTGTTGTCTGGAGGGAAGGACTGCTAGTTC	138		
QY	853	TGATTTTACTCTGTGAACACTTTATTAAAGCACTTCTTTTATTTGGCGGCTCTGTGA	912		
Db	137	TGATTTTACTCTGTGAACACTTTATTAAAGCACTTCTTTTATTTGGCGGCTCTGTGA	78		
QY	913	CCCTAGCCGCTTGACCCCGCTCTCTGTTGTACACTTTCAAGCAACACTTTTTCAGACTA	972		
Db	77	CCCTAGCCGCTTGACCCCGCTCTCTGTTGTACACTTTCAAGCAACACTTTTTCAGACTA	18		
QY	973	AAGGCCAAACAAAAGCT	989		
Db	17	AAGGCCAAACAAAAGCT	1		
RESULT 15					
LOCUS	BM119847/c				
DEFINITION	L0931C10-3 NIA Mouse Newborn Kidney cDNA Library (Long) Mus musculus cDNA clone NIA:L0931C10 IMAGE:30003009 3', mRNA sequence	727 bp	linear EST:09-JUN-2000		
ACCESSION	BM119847				
VERSION	EST.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus				
AUTHORS	Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.				
TITLE	Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method				
JOURNAL	Genome Res. 11 (9), 1553-1558 (2001)				
MEDLINE	21423098				
PUBMED	11544199				
COMMENT	On Nov 26, 2001 this sequence version replaced gi:17087873. Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@lgsun.grc.nia.nih.gov Plate: L0931 row: C column: 10 Seq primer: -21M13 Forward High quality sequence stop: 727 POLYA=Yes.				
FEATURES	Location/Qualifiers				
source	1..727 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="ntaEST:L0931C10-3" /db_xref="taxon:10090" /clone="NIA:L0931C10 IMAGE:30003009"				

/tissue_type="Newborn Kidney"
 /dev_stage="Newborn"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long)"
 /note="Organ: kidney; Vector: pSPORT1 (Invitrogen);
 Site.1: SalI; Site.2: NotI; Mouse cDNA project by the
 Laboratory of Genetics, National Institute on Aging (NIA),
 Intramural Research Program, NIH
 (http://agsun.grc.nia.nih.gov/cDNA/). This is a
 long-transcript enriched cDNA library (Ref. Genome Res.
 11:1553-1558 (2001). [PMID:11544199]). In
 brief,double-stranded cDNAs were synthesized with an
 oligo(dT) primer [Invitrogen: 5'-
 pGCACTAGTTCGATCGAGCGCGGCCCTTTT-3'] from 26
 microgram of total RNA, treated with T4 DNA polymerase,
 and purified by ethanol-precipitation. The cDNAs were
 ligated to lone-linker Li-Sal4, purified by
 phenol/chloroform, and separated from free linkers by
 Centricion 100. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-5. The products were purified
 by phenol/chloroform and Centricion 100. The cDNAs were
 digested with SalI and NotI enzymes, and cloned into
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
 coli host was transformed with ligation mixture by the
 standard chemical method. The average insert size is about
 3.0 kb. The library was constructed by Yulan Piao (NIA).

ORIGIN

Query Match	16.0%	Score 613.8	DB 12	Length 727
Best Local Similarity	99.7%	Prod. No. 1.6e-92		
Matches 615	Conservative 0	Mismatches 2	Indels 0	Gaps 0
373	AGCTGCTGGATCAGTGCCTGGCGGCACAGACACTTCTGGGGATGGTACAGGGTGTGACAA	432		
617	AGCTGCTGGATCAGTGCCTGGCGGCACAGACACTTCTGGGGATGGTACAGGGTGTGACAA	558		
433	TGACGTGCATCATCTTTGTTTCAAGCCCGAAACACAGTACAGCTTCAGGCACAGAGTG	492		
557	TGACGTGCATCATCTTTGCTTCAAGCCCGAAACACAGTACAGCTTCAGGCACAGAGTG	498		
493	GCAAGAGGAAACTGGAGGAGCACTGTCCACGGAGGGGCTCAAGACACCCGGCAACAGTG	552		
497	GCAAGAGGAACTGGAGGAGCACTGTCCACGGAGGGGCTCAAGACACCCGGCAACAGTG	438		
553	ACAAAAGAGCCCAAGAGGAGTATGCGTCAACCGGACCCCTGCCCATGTGGACTGTTTT	612		
437	ACAAAAGAGCCCAAGAGGAGTATGCGTCAACCGGACCCCTGCCCATGTGGACTGTTTT	378		
613	CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCTTAGCAGTGGGTAT	672		
377	CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCTTAGCAGTGGGTAT	318		
673	GAGGTGTGACGGGGCTGGGTGGCTTCTCCTCAGCCCATTAACAAGAGGGCCCCCACCCC	732		
317	GAGGTGTGACGGGGCTGGGTGGCTTCTCCTCAGCCCATTAACAAGAGGGCCCCCACCCC	258		
733	CCCCACGGCGAGCCTGGGAGGCTCTGCTGTCTCTTTAAGCCTCCTTACTCTCTTGGC	792		
257	CCCCACGGCGAGCCTGGGAGGCTCTGCTGTCTCTTTAAGCCTCCTTACTCTCTTGGC	198		
793	TCATCGACTATCGGTTCTGTGCTGTCTGTCTGTGTGGAGGAAAGCATGTGTATTC	852		
197	TCATCGACTATCGGTTCTGTGCTGTCTGTGTGTGGAGGAAAGCATGTGTATTC	138		
853	TGATTTTTACTCTGTGACACTTTATTTAAGACACTCTTTTTTATTCGGGCTCTGTGA	912		
137	TGATTTTTACTCTGTGACACTTTATTTAAGACACTCTTTTTTATTCGGGCTCTGTGA	78		
913	CCCTTAGCCGTTGACCCCGTCTCTGTGTGTA	972		
77	CCCTTAGCCGTTGACCCCGTCTCTGTGTGTA	18		